RICE UNIVERSITY

Development of Dynamic DNA Probes for High-Content in situ Proteomic Analyses

by

Ryan M. Schweller

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Michael R. Diehl, Chair
Assistant Professor, Bioengineering and Chemistry

Amina A. Qutub
Assistant Professor, Bioengineering

Mary C. Farach-Carson
Professor, Biochemistry & Cell Biology
Adjunct Professor, Bioengineering

HOUSTON, TEXAS
April 2012

ABSTRACT

Development of Dynamic DNA Probes for High-Content in situ Proteomic Analyses

by

Ryan M. Schweller

Dynamic DNA complexes are able to undergo multiple hybridization and dissociation events through a process called strand displacement. This unique property has facilitated the creation of programmable molecular detection systems and chemical logic gates encoded by nucleotide sequence. This work examines whether the ability to selective exchange oligonucleotides among different thermodynamically-stable DNA complexes can be harnessed to create a new class of imaging probes that permit fluorescent reporters to be sequentially activated ("turned on") and erased ("turned off"). Here, dynamic DNA complexes detect a specific DNA-conjugated antibody and undergo strand displacement to liberate a quencher strand and activate a fluorescent reporter. Subsequently, incubation with an erasing complex allows the fluorophore to be stripped from the target strand, quenched, and washed away. This simple capability therefore allows the same fluorescent dyes to be used multiple times to detect different markers within the same sample via sequential rounds of fluorescence imaging.

We evaluated and optimized several DNA complex designs to function efficiently for *in situ* molecular analyses. We also applied our DNA probes to immunofluorescence imaging using DNA-conjugated antibodies and demonstrated the ability to at least double the number of detectable markers on a single sample. Finally, the probe complexes were

reconfigured to act as AND-gates for the detection of co-localized proteins. Given the ability to visualize large numbers of cellular markers using dynamic DNA probe complexes, high-content proteomic analyses can be performed on a single sample, enhancing the power of fluorescence imaging techniques. Furthermore, dynamic DNA complexes offer new avenues to incorporate DNA-based computations and logic for *in situ* molecular imaging and analyses.

Acknowledgments

I would first like to thank my advisor, Dr. Michael Diehl for his guidance during my graduate career. He has taught me how to critically evaluate and perform research and for that I am extremely grateful. I would also like to thank my committee members, Dr. Amina Qutub and Dr. Cindy Farach-Carson, for critical reading and editing of this thesis.

Next, I have been blessed to work so closely with such a wonderful research group, Dr. Pamela Constantinou, Dr. Dzifa Duose, Dr. Arthur Rogers, Dr. Kenneth Jameson, Dr. Jonathan Driver, Matthew Zimmerman, Jan Zimak, Dr. Anand Radhakrishan, Dr. Artem Yefremov, Karthik Uppulury, Rocco Lucero, and Edward Samson. All of them have been great friends throughout my PhD and many have contributed directly to the work in this thesis. Similarly, several undergraduate researchers deserve thanks for their contributions to my research, including Nick Frankel, Jessica Joyce, Lisa Tseggay, and Harsha Mishra.

In addition, I would like to acknowledge all of our collaborators. Dr. Walter Hittelman has been a wonderful sounding board for our research efforts as well as critical evaluator of our publications and data. Similarly, Dr. Qutub has always offered input and methods to improve the quality and impact of our work.

I owe a great deal of thanks to all of my friends and family. The many friends I have made at Rice have been a tremendous help in all my endeavors. I would specifically like to thank my girlfriend Morgan Gallagher for all her support and patience.

Lastly, I would like to thank my parents, Mike and Lynne, and my brother, Eric, who have always encouraged me. I love you all.

Contents

List of Figures

List of Tables

List of Abbreviations

BLAST	Basic Logic Alignment Search Tool
CaM	Calmodulin
CSC	Cancer Stem Cell
DNA	Deoxyribonucleic Acid
dsDNA	Double Stranded DNA
EDC	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-Linked Immunosorbent Assay
EMCCD	Electron-Multiplying Charge-Coupled Device
FBS	Fetal Bovine Serum
FISH	Fluorescence *in situ* Hybridization
FPLC	Fast Protein Liquid Chromatography
GFP	Green Fluorescent Protein
GST	Glutathione S-Transferase
HCR	Hybridization Chain Reaction
HPLC	High-Performance Liquid Chromatography
IF	Immunofluorescence
KLC4	Kinesin Light Chain 4
MALDI-TOF	Matrix-Assisted Laser Desorption/Ionization – Time-of-Flight
MELC	Multi-Epitope-Ligand Cartography

mRNA	Messenger RNA
NHS	N-Hydroxysuccinimide
PAGE	Poly-Acrylamide Gel Electrophoresis
PCR	Polymerase Chain Reaction
PLA	Proximity Ligation Assay
RCA	Rolling Circle Amplification
RNA	Ribonucleic Acid
SNR	Signal-to-Noise Ratio
sulfo-SMCC	Sulfosuccinimidyl 4-[*N*-maleimidomethyl]cyclohexane-1-carboxylate
ssDNA	Single Stranded DNA
TAE	Tris-Acetate-EDTA
TCEP	tris(2-carboxyethyl)phosphine
V/V	Volume per volume
WASP	Wiskott-Aldrich Syndrome Protein
W/V	Weight per volume

Chapter 1

Introduction

1.1. The Need for Fluorescence Imaging Techniques

Bulk protein detection techniques have become important to proteomics research due to their ability to detect and analyze large numbers of markers in parallel, including two-dimensional gel electrophoresis [1], western blotting [2], mass spectrometry [3], antibody microarrays [4], and enzyme-linked immunosorbent assays (ELISAs) [5]. However, these methods typically require destructive preparation of samples to create homogeneous analyte solutions. Such processing results in the loss of spatially-encoded marker levels and distributions. For example, solid tumors are composed of a combination of benign, cancerous, and stromal cells, and consequently, techniques that analyze these architectural features are central to our understanding of tumor progression and diagnosis. Furthermore, the ability to monitor biomarkers in intact cells and tissues is also important for many analyses because individual cell phenotypes can be distinguished due to their external environment and boundary conditions [6]. For these

reasons, fluorescence microscopy approaches have therefore become critical to biomarker detection and screening for the characterization of healthy and diseased tissues.

1.2. Current Limitations in Fluorescence Imaging

Although fluorescence detection methods have broad utility for *in situ* proteomic analyses, fluorescence microscopy techniques are largely unable to dissect levels or spatial distributions of markers in cell or tissue samples. This defficiency is primarily due to the inability to spectrally separate large numbers of fluorescent dyes. While improvements have been made through instrumentation (i.e. confocal or hyperspectral microscopy) [7-11], fluorescence imaging is incapable of capturing the full cellular 'picture' needed for proteomic analyses. The standard solution for this problem is to assess small numbers of markers on large numbers of samples to generate statistical correlations among target proteins. These advances are effective in some analyses. However, they are prohibitive when samples are highly heterogeneous or when rare cell types are to be interrogated. For example, cancerous tissues exhibit high heterogeneity among cell morphologies and protein expression patterns that can have very important functional significance (Figure 1.1) [12-14]. Yet, the inability to detect more than a few dyes makes it difficult to assess functionally important relationships among multiple samples [15]. This problem has greatly hindered the identification of reliable cancer biomarkers for prognosis. While many biomarkers have been discovered to date, far fewer have been integrated into clinically relevant assays due to the lack of large scale functional correlations among individual markers [16-18].

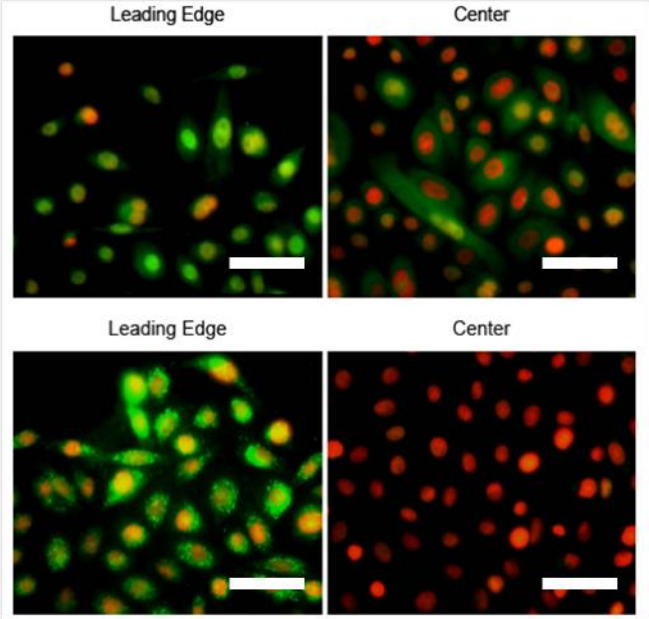

Figure 1.1 Spot cultures of two immortilized bronchial epithelial tumor cell lines (top and bottom). The cells are stained for phosphorylated EGFR (green) and nuclei (red). In the top images, cells show similar phosphorylation in the center of the culture as well as at the edge of the culture, but with large heterogeneity within the images. The bottom cells show high heterogeneity between the leading edge and center of the culture, but a very homogeneous response within the images. Adapted from [19]. (Scale bars: 100 μm)

The inability to detect more than a few markers make analyses of rare cell types impossible as the entire suite of fluorescent markers is often needed to identify the relevant cells. For example, cancer stem cells (CSCs) often represent < 1% of a tissue's cell population (Figure 1.2) [20] and reproducibility of detection is hotly debated [21]. Although a vast number of stem cell markers are known for a variety of cells and tissues [22-28], it has proven difficult to isolate CSCs and devise more effective methods of detection. These cells are particularly important as they have demonstrated a high resistance to most standard cancer treatments, including radiation therapy [29, 30],

potentially enriching CSC populations [20]. Furthermore, CSCs exhibit 'stem-like' behavior, allowing them to recapitulate an entire tumor on their own [21], making effective identification and treatment a high priority.

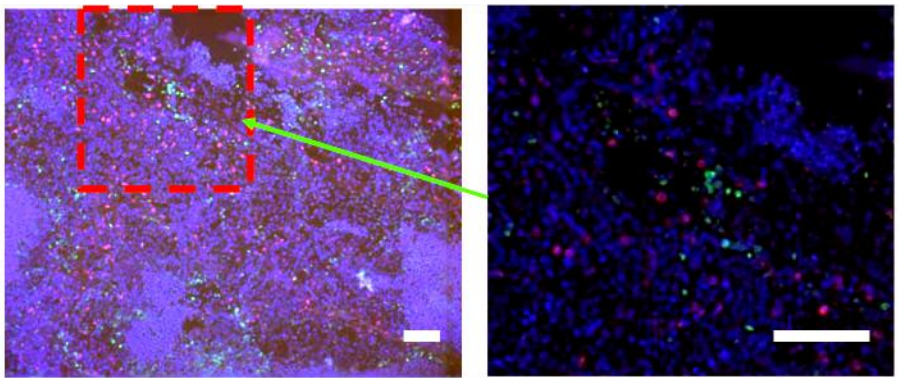

Figure 1.2 Tissue section of a H460 lung tumor xenograft. Animals were treated with IdUrd (green) and ClrdUrd (red) and cells were stained with DAPI (blue). The green cells are believed to be 'dormant' CSCs, while red cells are believed to be proliferating progenitor stem cells. Adapted from [19]. (Scale bars:250 μm)

1.3. Current Fluorescence Imaging Technologies

Immunofluorescence (IF) is a common technique for cellular analyses, harnessing fluorescently labeled antibodies for detection. As antibodies can be raised against nearly any biological target (i.e. proteins, haptens, and nucleic acids) with high specificity and affinity, they are ideal targeting agents for a variety of fluorescence-based immunoassays. Both direct and indirect IF have been mainstays for cellular imaging (Figure 1.3) and IF capabilities have been continually expanded by developing enzymatic amplification methods [31], employing a variety of different fluorescent reporter molecules such as quantum dots [14, 32], and devising methods to utilize primary antibodies form the same

host species [33, 34]. As an alternative to immunofluorescence, DNA-based methods have also shown great promise through hybridization-based detection assays.

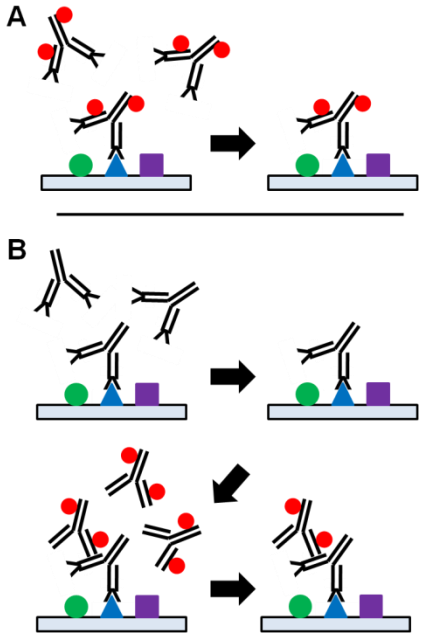

Figure 1.3 Schematics of direct (A) and indirect immunofluorescence (B).

1.4. DNA-Based Fluorescence Imaging Techniques

DNA has become increasingly attractive for imaging applications because of the ability to generate large numbers of orthogonal nucleotide sequences (i.e. 4^n nucleic acid combinations) and incorporate diverse sets of fluorophores or functional groups. Fluorescence *in situ* hybridization (FISH) is commonly used to detect and even count specific chromosomal DNA sequences and mRNA transcripts [35-40] as well as track their transport in live cells [41, 42]. This technique has granted a better understanding as to how cells regulate protein synthesis with high spatio-temporal control [43]. Molecular

beacons are a special class of FISH probe that take on a stem-loop, hairpin conformation. The ends of the stem contain dye and quencher molecules, so fluorescence is only produced upon hybridization and separation of the stem [44-48].

DNA-based methods have also been developed for the detection of proteins. DNA and RNA aptamers have offered the most direct methods to fluorescently label protein markers in cellular environments [49-51]. DNA also can be directly conjugated to antibodies for fluorescence detection applications. Originally, these conjugates were used to label antigens with ssDNA, which could then be detected and quantified via polymerase chain reaction (PCR) [52]. Since this initial work, *in situ* imaging techniques have been developed employing rolling circle amplification (RCA) [53, 54]. RCA extends the DNA label into a long repeating polymer, which can be detected via the hybridization of short fluorescent oligonucleotides. This work has been expanded to also detect protein-protein interactions through proximity ligation assay (PLA) [55, 56], where two DNA-labeled antibodies must be close in close proximity to form the template for RCA. Recently, Teo *et al.* synthesized a class of fluorescent DNA analogues capable of being excited by UV light [57]. By combining these synthetic bases, unique polyfluorophores-based DNA backbones have been created to label antibodies for IF [58], similarly to standard fluorescent dyes.

Although DNA offers new fluorescence imaging modalities and a method to generate large numbers of orthogonal targets for proteomic analyses, it still cannot overcome the basic spectral limitations of fluorescence imaging. As such, dynamic DNA technologies, which are capable of undergoing multiple hybridization and dissociation

events, may offer an alternative route to addressing the limitations of fluorescence imaging.

1.5. Dynamic-DNA for Fluorescence Imaging Applications

While static DNA technologies terminate upon the formation of a thermodynamically stable conformation, dynamic DNA interactions are capable of function in complex non-equilibrium reactions, acting as logic gates [59, 60], motors [61-64], and computational networks [65-67]. Dynamic DNA reactions employ a very specialized and selective DNA interaction known as branch migration. Here, a ssDNA hybridizes a short 'toehold' domain (**1**, Figure 1.4A) and branch migrates, displacing a strand from the original DNA complex (**2**, Figure 1.4A). This process can be expanded by incorporating a toehold dissociation step to reveal a new toehold domain (**3**, Figure 1.4B).

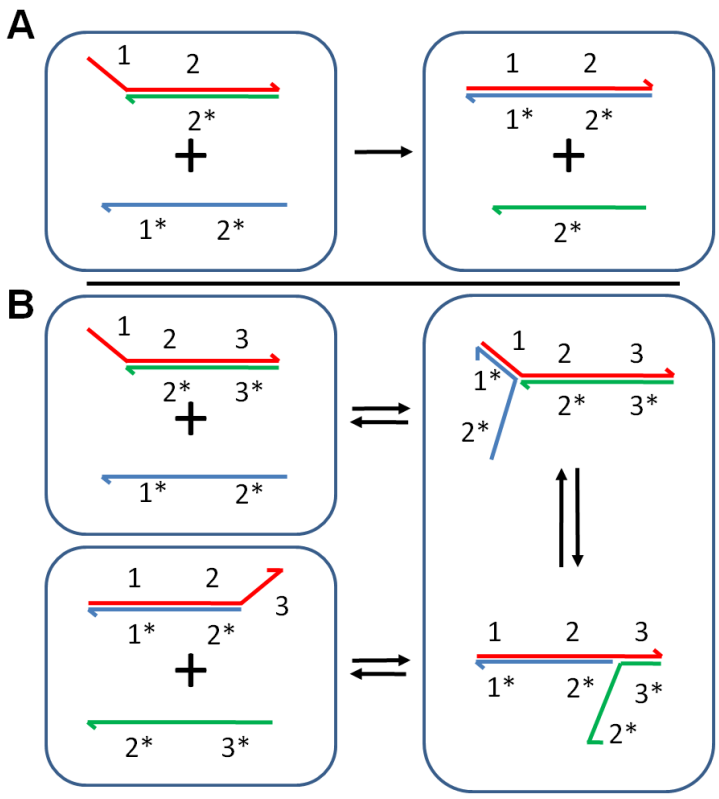

Figure 1.4 Mechanisms of three-way branch migration-based strand exchange reactions. A. Toehold-mediated strand displacement, and B. Toehold mediated strand exchange.

By treating these reaction mechanisms as simple building blocks, increasingly complex reaction networks have been created. Zhang *et al*. designed a catalytic 'circuit' that combines two strand exchange steps, creating a modular cyclic reaction network. They could then combine multiple reconfigured 'circuits' to exhibit different catalytic behaviors (i.e. cross-catalysis and autocatalysis) or act as an analog AND-gate [68]. Similarly, Yin *et al.* devised a basic input-output method to construct biomolecular networks based upon simple DNA hairpin motifs [63]. By controlling the arrangement and display of toehold domains on these hairpins, they can create catalytic circuits,

arbitrary dendrimer structures, or walkers, illustrating the versatility of biomolecular functions that can be encoded in DNA. Furthermore, dynamic DNA behaviors have also been used for fluorescence imaging. Seferos *et al.* illustrate the first direct incorporation of strand exchange for cellular detection [69]. In this work, ssDNA was hybridized to a short fluorophore-bearing oligonucleotide and conjugated to a gold nanoparticle, quenching the fluorescence. Upon binding a target mRNA transcript the fluorophore strand is released via branch migration, allowing the fluorescence to be visualized. Since this initial work, "nano-flares" have been multiplexed to detect multiple mRNA transcripts simultaneously using a single nanoparticle [70].

More recently, Choi *et al*. employed a hairpin-based amplification scheme, known as hybridization chain reaction (HCR) [71], to fluorescently localize five distinct mRNA transcripts in fixed zebrafish embryos [72]. The hairpins are capable of "opening" upon the detection of a specific mRNA transcript. The previously looped DNA sequence then acts as a new toehold allowing a second hairpin to bind and open. This process continues similar to a chain reaction, requiring only two hairpin motifs to create a sequential linear polymer structure. By incorporating dye molecules into the hairpin structures, the DNA polymer acts a linear signal amplifier.

1.6. Objectives and Summary

This work seeks to develop a novel fluorescence labeling scheme, where the same fluorescent dyes are used multiple times to detect different markers within a single cell or tissue sample via multiple rounds of fluorescence imaging. To accomplish this, reactive DNA complexes were designed that undergo strand displacement reactions to

sequentially activate a fluorescent signal upon the detection of a target DNA sequence and subsequently deactivate the signal upon the addition of an erasing complex. In this thesis, we will create DNA-conjugated antibodies capable of detecting target antigens, design and optimize DNA complexes to facilitate the *in situ* detection of protein markers, selectively label and image protein markers on fixed cells using DNA-conjugated antibodies, and utilize the reconfigurable nature of dynamic DNA complexes to implement logic-based detection schemes.

Several other techniques have been developed to facilitate multiple rounds of fluorescence imaging for the detection of large numbers of markers. Schubert *et al.* developed a method known as multi-epitope-ligand cartography (MELC), allowing them to detect and analyze 49 molecular markers *in situ* [73]. Although this technology represents a great leap forward in fluorescence imaging for proteomic analyses, it is not readily adaptable to a laboratory setting. Fifteen automated work stations were required to generate their data sets and only two fluorophores could be visualized simultaneously. Furthermore, deactivation of fluorescent reporters was done via photobleaching and long term exposure to UV light sources could be damaging to cell or tissue samples. Similarly, Suzuki *et al.*, developed a heat treatment method to eliminate the antigenicity of primary antibodies [74]. By heating samples to 90º C for fifteen minutes, they found primary antibodies could be removed from their sample, allowing primary antibodies from the same species to be used for subsequent labeling. Yet, they do not test their detection ability beyond three protein markers and the effects of multiple heat cycles on a sample are unknown.

Employing reactive DNA complexes to activate and erase fluorescent signals, all desired antibodies could be incubated on the sample simultaneously and detected in small sets at room temperature, under mild buffering conditions (generally TAE with magnesium or PBS). In addition, activating and erasing reactions can be performed simultaneously, minimizing wait times between fluorescence imaging sessions. Lastly, the strand displacement mechanisms used to activate and erase the fluorescent signals are the most basic reactions employed by dynamic DNA technologies. Therefore, while heat disruption and photobleaching may be amenable to "erase" IF and DNA-based fluorescence detection, these methods cannot recapitulate the diverse capabilities offered by strand displacement mechanisms to create chemical circuits or computational networks for new and unique fluorescence detection modalities.

To illustrate the development and use of our dynamic DNA probe complexes, this thesis is arranged as follows:

Chapter 2 will discuss the original design of our reactive DNA probe complexes using the entropic-driven circuit components developed by Zhang *et al.* [68]. We will then test these probes on a microarray platform as well as on fixed cells as a proof-of-principle demonstration for this technology.

In Chapter 3, we will optimize the DNA probe complexes to operate efficiently in fixed-cell environments. While most dynamic DNA reactions are performed in homogeneously mixed solution-phase environments, the DNA probes should be tuned to reflect the cellular reaction environment.

Chapter 4 will examine the direct application of the DNA probe designs to immunofluorescence imaging. Employing antibody-based detection, the probe

complexes should be able to exhibit responses similar to those seen by fluorophore-labeled secondary antibodies as well as allow multiple rounds of fluorescence microscopy.

Chapter 5 will demonstrate the ability to apply more complex detection methods using dynamic DNA probes. By modifying our probe complexes, we will build an AND-gate capable of identifying co-localized proteins on cell samples.

Finally, in Chapter 6, we will conclude these studies and detail future avenues and experiments for this work. Chapter 7 is purely supplemental, detailing the synthesis and characterization of our semi-synthetic DNA-protein conjugate used for studies in Chapters 2, 3, and 5. It has been included as a chapter to maintain formatting and properly incorporate references.

Chapter 2

Multiplexed and Reiterative Fluorescence Labeling via DNA Circuitry

This chapter describes the design of fluorescent DNA-based probes for fluorescence imaging. We hypothesized that the half reactions from entropically driven circuits developed by the Zhang *et al.* (ref [68]) could be used to selectively label and erase protein markers within fixed cells. These circuits are composed of three components: a substrate complex, catalyst strand, and fuel strand. Normally, all three components are present to drive the reaction forward converting substrate to waste and regenerating the catalyst. By incorporating quencher and dye molecules into the substrate complex and fuel strand, a fluorescently active intermediate complex can be trapped kinetically. After imaging, the fuel strand can be added to quench the fluorophore and regenerate the catalyst.

To achieve this, three metrics need to be accomplished: 1. Reactions should label catalysts specifically, 2. Circuits should be capable of operating in parallel, and 3. Catalyst (target) strands should be capable of being multiple several times. To test this, we developed a microarray platform to test and evaluate probe complexes, including their reaction efficiencies and cross-reactivity. This platform then was used to demonstrate parallel operation of circuits. Furthermore, using the microarray platform we show that we are able to detect four unique target strands with only two fluorescent dyes. Finally, using a transfectable GFP fusion protein, which can be labeled using an engineered DNA-protein conjugate, we demonstrate selective fluorescence activation as well as the ability to label a single target strand multiple times. This chapter, thus, acts as a proof of principle for reiterative and reiterative DNA-based detection of cellular markers.

Reproduced with permission from:

Duose, Dzifa Y.;* Schweller, Ryan M.;* Hittelman, Walter N.; Diehl, Michael R.; Multiplexed and Reiterative Fluorescence Labeling via DNA Circuitry. *Bioconjugate Chem.*, 2010, 21 (12), 2327–2331.

*Equal Contributors

2.1. Introduction

In situ imaging of molecular markers has become increasingly important to advancements in the biological and biomedical sciences [35, 38, 75-77]. These approaches offer advantages over bulk analyses, since they allow spatially dependent

expression patterns of RNA and proteins to be delineated in cells and tissues. Several antibody and nucleic acid-based fluorescent probe technologies have been developed for marker imaging, and the sensitivity of these probes has been continually improved [45, 77-79]. Yet, many contemporary cytological studies now require increasingly comprehensive molecular pathway analyses to characterize the network-level properties of cells and resolve functional relationships among cell phenotypes and their tissue distribution [32, 80, 81]. In such cases, the number of markers one seeks to examine can easily exceed the number of probes that can be used simultaneously for detection due to the spectral overlap of the reporting dye molecules. Thus, various biological studies stand to benefit from methods that allow a greater number of molecular markers to be visualized.

The number of markers that can be evaluated on an individual biological sample could be increased if it were possible to remove fluorescent probes from cells such that new markers could be labeled and detected using the same fluorescent reporting molecules. However, dye molecules are typically attached covalently to probes that are engineered to specifically bind their targets with high affinity. Probe removal therefore requires the use of harsh chemical and/or physical treatments that can disrupt cell and tissue morphology and compromise subsequent marker analyses.

To overcome these limitations, we adapted a new class of DNA circuit technologies as molecular imaging probes that allow fluorophores to be sequentially coupled to, and removed from, molecular markers using exceptionally mild processing conditions. Derived from “entropically driven” circuits developed by Zhang et al. [68], these probes operate using principles distinct from those of FISH probes and molecular

beacons. Whereas these latter technologies rely on more classical, sequence-dependent hybridization reaction mechanisms for target recognition, the different complexes that make up the entropically driven circuits are designed to react with one another via a process called strand-displacement: the selective exchange of individual oligonucleotides among different complexes of DNA [59, 63, 82, 83]. These reactions proceed in a sequence-dependent fashion, and hence, multiple circuits can be employed simultaneously to label different targets. Yet, the total number of matched base pairs that are formed through these reactions does not necessarily have to increase for them to proceed efficiently (i.e., reactant and product complexes can be isoenergetic). As a result, strand displacement reactions can be used to assemble, isolate, and disassemble stable intermediate-state complexes of multistep reaction cycles. Herein, we show that such capabilities can be harnessed to create erasable molecular imaging probes that can function at ambient temperatures and in mild, nondenaturing buffers (e.g., Tris-based buffers). Thus, while the circuit-based probes facilitate multiplexed (simultaneous multicolor) detection, they should also allow the same-colored dye molecules to be used reiteratively to label different markers on an individual sample of fixed cells.

2.2. Materials and Methods

2.2.1. General

Oligonucleotides were purchased from Integrated DNA Technologies (IDT Coralville, IA) and can be found in Table 2.1. The recombinant target protein GFP-Z_E, was designed with C-terminal leucine zipper (Z_E) and 6× Histidine tag (6xHis) for

purification using standard cloning procedures. Artificial proteins were labeled with catalyst DNA as described in ref [84].

2.2.2. DNA Circuit Design and Characterization

The DNA circuit sequences were designed as previously described [68]. Fluorophores and quenchers were incorporated into substrate complexes as shown in Figure 2.1. Strands outfitted with fluorophores or quenchers were purchased HPLC-purified. DNA complexes were formed via a thermal annealing procedure: strands were mixed together at a 1:1 stoichiometry in TAE/Mg^{2+} buffer at a final concentration of 3 μM. The temperature of this solution was then raised to 95° C and reduced to 25° C over 90 min. DNA complex formation was verified by 12% (w/v) non-denaturing PAGE gel analyses using SYBR-Gold staining (Invitrogen Carlsbad, CA).

Figure 2.1 Circuit design scheme with strand labeling convention for complexes.

Sequence Name	Nucleic Acid Sequence
Circuit1_O1	CCA CAT ACA TCA TAT TCC CTC ATT CAA TAC CCT ACG/3IAbRQSp/
Circuit1_O2	CTT TCC TAC A CC TAC GTC TCC AAC TAA CTT ACG G
Circuit1_LB	TGG AGA /iCy5/CGT AGG GTA TTG AAT GAG GGC CGT AAG TTA GTT GGA GAC GTA GG
Circuit1_C	CAT TCA ATA CCC TAC GTC TCC ATT TTT TTT TT /3AmMC6/
Circuit1_Fuel	CCT ACG TCT CCA ACT AAC TTA CGG CCC TCA TTC AAT ACC CTA CG/3IAbRQSp/
Circuit1_O1 comp	GTA TTG AAT GAG GGA ATA TGA TGT ATG TGG
Circuit1_O1 short	CCA CAT ACA TCA TAT TCC CTC ATT
Circuit2_O1	ACC TCT TCA CGA ACA TTT CA/3IAbRQSp/
Circuit2_O2	ACC TAA TAG C AC CAC ATC AAT CTC GAT CCA GTA C
Circuit2_LB	TGG CTA /iCy3/TGA AAT GTT CGT GAA GAG GTG TAC TGG ATC GAG ATT GAT GTG GT
Circuit2_C	CTT CAC GAA CAT TTC ATA GCC ATT TTT TTT TT /3AmMC6/
Circuit2_Fuel	ACC ACA TCA ATC TCG ATC CAG TAC ACC TCT TCA CGA ACA TTT CA
Circuit2_O1 comp	TGT TCG TGA AGA GGT
Circuit2_O1 short	ACC TCT TCA
Circuit3_O1	TCA CAC ATC AAC CTC T TCTT T CTC TCG ACA CAT CAC
Circuit3_O2	CTT TCC TAC A CT TAT TCA TCC TTT CAC TCA CTT C
Circuit3_LB	GAA GTG AGT GAA AGG ATG AAT AAG AAG AGTG ATG TGT CGA GAG AAAG TAA
Circuit3_C	TCT CTC GAC ACA TCA C TTA CTT TT TTT TTT TT /3AmMC6/
Circuit3_Fuel	CT TAT TCA TCC TTT CAC TCA CTT CTCTT TCT CTC GAC ACA TCA C

Table 2.1 List of nucleic acid sequences used for conjugation, microarray experiments, and multiprotein scaffold experiments. /3AmMC/, /3IAbRQSp/, /iCy5/, /iCy3/, /5Cy5/, and /5Cy3/ indicate 3′ amine, 3′ Iowa Black quencher, internal Cy5, internal Cy3, 5′ Cy5, and 5′ Cy3 modifications, repectively.

Sequence Name	Nucleic Acid Sequence
Circuit3_O1 comp	TGT CGA GAG AAA GAA GAG GTT GAT GTG TGA
Circuit3_O1 short	TCA CAC ATC AAC CTC TTC TTT CTC
Circuit4_O1	CAC CAA CCC AAT TCT C CCTA C CCA TTC CTG TAT CAT
Circuit4_O2	ACC TAA TAG C TAC CTT CCC TCT ATT CAT GTC CAC
Circuit4_LB	GTG GAC ATG AAT AGA GGG AAG GTATAG GATG ATA CAG GAA TGG GTGG AGT
Circuit4_C	CCC ATT CCT GTA TCA T AC TCC A TT TTT TTT TT /3AmMC6/
Circuit4_Fuel	TAC CTT CCC TCT ATT CAT GTC CAC CCT ACC CAT TCC TGT ATC AT
Circuit4_O1 comp	CAG GAA TGG GTA GGG AGA ATT GGG TTG GTG
Circuit4_O1 short	CAC CAA CCC AAT TCT CCC TAC CCA
Circuit5_O1	CAT ACC ACA ACA ATT TAC TTC ACC AAC CCA TCC ACT
Circuit5_O2	CTT TCC TAC AAA TCG CCA AAC TAC AAA CTC AAT C
Circuit5_LB	GAG GTG AGT GGA TGG GTT GGT GAA GTG ATT GAG TTT GTA GTT TGG CGA TT
Circuit5_C	CAC CAA CCC ATC CAC TCA CCT C TT TTT TTT TT /3AmMC6/
Circuit5_Fuel	AAT CGC CAA ACT ACA AAC TCA ATC ACTT CAC CAA CCC ATC CAC T
Circuit5_O1 comp	TGG GTT GGT GAA GTA AAT TGT TGT GGT ATG
Circuit5_O1 short	CAT ACC ACA ACA ATT TAC TTC ACC
1compcy5	/5Cy5/TGT AGG AAA G
1compcy3	/5Cy3/GCT ATT AGG T

Table 2.1 Cont.

2.2.3. Microarray Procedures

The DNA microarrays were printed on Vantage silyated aldehyde slides (CEL Associates Pearland, TX) using SMP3 pins (ArrayIt Sunnyvale, CA) and a custom

fabricated microcontact printer. Arrays were fabricated by spotting solutions of 3′-amine labeled C strands (10, 5, 1, 0.5, 0.1, and 0.05 μM stocks for gradient experiments and 2 μM stocks for all others) in PBS (pH 6.6) containing 30% (w/v) glycerol. Afterward, the slides were incubated in a humidity chamber for 6 h. Free aldehyde groups on the slides were then quenched for 5 min in a sodium borohydride solution [3:1 PBS/EtOH 2.5% (w/v) $NaBH_4$]. Slides were blocked for 2 h in 4× SSPE buffer (600 mM NaCl, 40 mM NaH_2PO_4, and 5 mM EDTA) containing 0.1% (w/v) BSA, washed 3 times with 4× SSPE buffer with 0.1% (w/v) SDS, rinsed with milliH_2O, and dried under nitrogen.

Microarray labeling/activation and dye removal/deactivation reactions were performed using a static incubation procedure or with a hybridization station (TrayMix2: ArrayIt) that provides active mixing of reagents over the slide. For static incubation, Gene frames (AbGene Rockford, IL) were affixed over the arrays to create a reaction chamber containing 5 pmol substrate and OC1 consumption complex in TAE w/Mg^{2+}. The arrays were incubated overnight. The Gene frames were then removed and slides were washed 3 times in 4× SSPE buffer, rinsed again in milli-H2O, and dried under a nitrogen stream. Deactivation and reiterative labeling experiments were performed similarly by affixing new Gene frames to the slides and repeating the incubation procedure. Microarray analysis was performed using ImageJ (http://rsbweb.nih.gov/ij/).

For Figure 2.3b, average fluorescent intensity was measured for each spot concentration (ten replicates of each concentration). Averages (+/- standard deviation) were plotted against the spotted concentration and fit to a Langmuir model. All statistical analyses were performed using OriginPro 8 (OriginLab, Northampton, MA).

2.2.4. Cell Culture and Labeling

HeLa cells were cultured in an 8-chambered coverslips (Lab-Tek Rockford, IL) for 24 h in DMEM supplemented with 10% (v/v) FBS, 50 µg/mL penicillin, and 50 µg/mL of streptomycin. For GFP labeling experiments, the media was replaced and the cells were transfected with GFP-Z_E DNA using Fugene (Roche Indianapolis, IN) transfection reagent under the manufacturer's protocol.

To label cells, coverslips were washed once with PBS and fixed with freshly prepared 4% (w/v) paraformaldehyde (PFA) solution for 30 min. The cells were then washed twice for 2 min with PBS, permeabilized for 5 min with 0.2% (v/v) Triton X-100, washed twice again with PBS, and stored overnight at 4 °C. Prior to circuit labeling experiments, the cells were washed again with PBS and then incubated for 2 h with a blocking solution [1% (w/v) BSA, 1 mg/mL denatured Herring sperm DNA, and 0.5 µM polyT DNA in PBS]. For GFP labeling experiments, the cells were also incubated with 400 nM of Z_R-ELS_6-Cat1 in PBS for 2 h. Excess polymer was then removed by washing twice with PBS prior to circuit-based labeling.

Circuit labeling reactions were carried out by incubating the cells for 1.5 h with 100 nM substrate complex and then washing twice with PBS for 2 min. Dye removal was performed similarly using 1 µM fuel. Before imaging experiments, slides were washed twice with PBS and then mounted on a glass slide using rubber cement.

2.2.5. Fluorescence Imaging and Analysis

All images were collected using a Zeiss Axioplan epifluorescence microscope and are contrasted identically in each figure. Correlations between GFP and Cy5-circuit

signals and Cy5 signals produced during sequential labeling of cells (ON1 and ON2) were analyzed using a custom program written in MATLAB.

2.3. Results and Discussion

An illustration depicting our use of entropically driven circuits as imaging probes is shown in Figure 2.2. Adopting the nomenclature of Zhang et al. [68], these circuits are composed of three main components: a single-stranded "catalyst" strand (C), a three-strand DNA complex called the "substrate" (S), and a "fuel" strand (F). In the circuit reaction cycle, the binding of C to a 6 nucleotide "toehold" domain at one end of S initiates a strand displacement reaction that releases an "output" strand (O1) and produces an intermediate-state complex (IR) that possesses a new, internal 4 bp toehold domain. The binding of F to this toehold then initiates a second strand displacement reaction that releases C from the IR complex and produces a "waste" product (W). To convert these complexes into imaging probes, catalyst strands can be appended to targeting agents that bind to specific molecular markers. The circuit substrates are modified by incorporating fluorescent dye (Cy3 or Cy5) and quencher (Iowa Black) molecules that are positioned such that the reaction of S with C results in an IR complex that contains an unquenched fluorophore. The dye-bearing strand within IR can then be removed from the marker and rendered inactive in the waste product by incubation with a modified F strand that carries a second quencher molecule. Overall, the use of quenched substrates and waste products should reduce background fluorescence resulting from potential nonspecific binding of either complex to a sample. Finally, we note that, in contrast to prior work where C truly functioned as a catalyst [68], marker labeling and removal in the present application is

achieved by performing partial circuit reactions while using C for targeting. We therefore use the term catalyst only for continuity with previous reports.

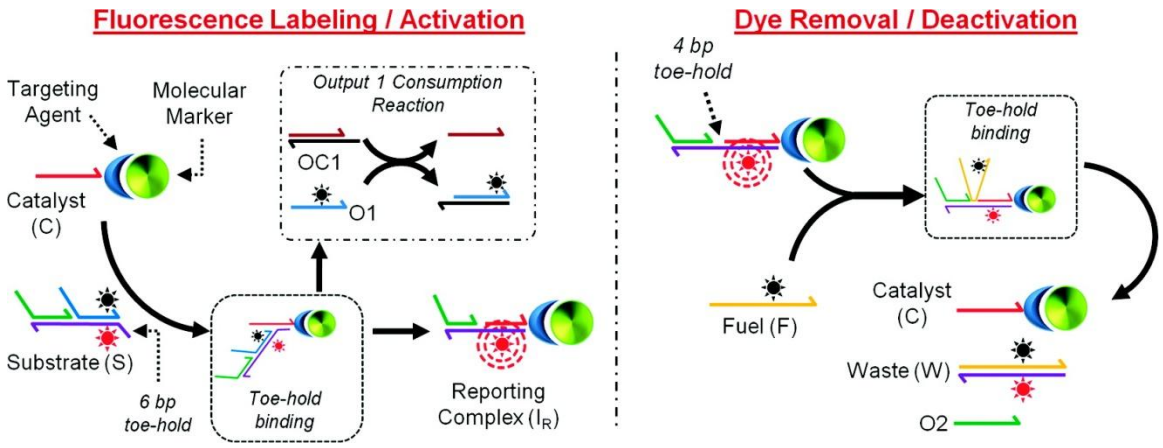

Figure 2.2 DNA Circuit-Based Marker Labeling and Dye Removal Reactions

To evaluate the efficiency of the circuit-based labeling and dye removal reactions, we first examined distributions of product complexes that were formed upon incubations of S and C, as well as S, C, and F via native PAGE-gel analyses (Figure 2.3a). After a partial circuit reaction of S with C, the catalyst strand is bound to the IR complex through a total of 22 matched base pairs. This complex is therefore stable at room temperature and can be isolated on a gel. Yet, the free energy difference between S and IR is small (ΔG ~0.4 kcal/mol). As a result, the reaction of S and C, when performed using equimolar concentrations, results in an equilibrium distribution of circuit components possessing near-equivalent concentrations of S, IR, and free, "unlabeled" catalyst (Figure 2.3a, lane 5). While this result is not optimal for marker labeling, this reaction can be driven forward by adding a second complex (OC1) that consumes (O1) once it is liberated from S (boxed reaction in Figure 2.2). As demonstrated in other strand

displacement systems [85], the sequestration of O1 shifts the equilibrium distribution of the reaction significantly toward the IR state (Figure 2.3a, lane 3). Alternatively, this distribution can also be shifted using an excess of S relative to C, which should often be the case when labeling catalysts (markers) that are immobilized on a specimen. Nevertheless, the ability to drive strand displacement reactions forward through output sequestration will likely be useful for optimizing dye labeling kinetics or when local target concentrations within a sample are high.

The removal of dyes from a sample via the reaction of IR and F constitutes an equally important step in our marker imaging procedure. Here, the use of three-strand S and IR complexes, as opposed to somewhat simpler two-strand complexes, allows dye-bearing strands to be displaced from the reporting IR complex without output sequestration (Figure 2.3a, lane 4), since two strands (C and O2) must react simultaneously with final circuit product W for the reverse reaction W + O2 + C → IR + F to occur. Hence, the dye removal reaction is effectively irreversible.

We next performed a series of DNA microarray experiments that were designed to evaluate the use of DNA circuit complexes as molecular imaging probes (Figure 2.3). In these experiments, amine-modified C strands are arrayed on the surfaces of glass slides, and the reaction of S with C produces a fluorescent IR complex that is anchored to the slide surface and can be detected using a fluorescence microarray scanner. Analyses of arrays where a catalyst was printed at variable spot concentrations confirm that the IR complex of the circuit can function as a reliable reporter of the levels of immobilized catalyst (Figure 2.3b, top). Here, spot intensity profiles can be approximated by the Langmuir adsorption equation, as is commonly found with DNA microarrays [86, 87].

Furthermore, after the same slide is incubated with a fuel strand, each spot disappears and cannot be detected over background autofluorescence signals of the slide (Figure 2.3b, bottom).

Our microarray experiments also allowed us to demonstrate the use of DNA circuit probes for both multiplexed and reiterative marker labeling. We found that immobilized C strands can be labeled with dyes multiple times via sequential reactions of arrays with S, F, and then a second solution of S complexes (Figure 2.3c). Each labeling reaction produced arrays possessing near identical spot intensities. We also demonstrated that multiple C targets could be labeled and/or erased simultaneously using multiple S and F complexes in a single reaction step (Figure 2.3d). In these experiments, five different C strands were printed both as mixtures of two strands and individually on the surface of the slide. The array was then reacted with two different substrate complexes (S1-Cy5 and S2-Cy3), yielding a spot pattern that corresponded directly to the positions of the printed catalysts (C1 and C2). Subsequently, in a single incubation step, this pattern was erased with F1 and F2, and a new spot pattern was generated through a reaction with a second set of substrates (S3-Cy5 and S4-Cy3); spotted lanes where C3 and C4 were printed appear in the scanned image. Throughout this procedure, the fifth printed catalyst strand of the array (C5) remains unlabeled in both scanned images. Thus, these experiments confirm that DNA circuitry can be used for multiplexed and reiterative imaging: two fluorescent dye molecules and two spectral channels of an imaging system are used to detect four distinct markers on the same sample. Importantly, all labeling and removal reactions in these assays were performed using mild processing conditions (room temperature and Tris buffer supplemented with 12.5 mM Mg^{2+}).

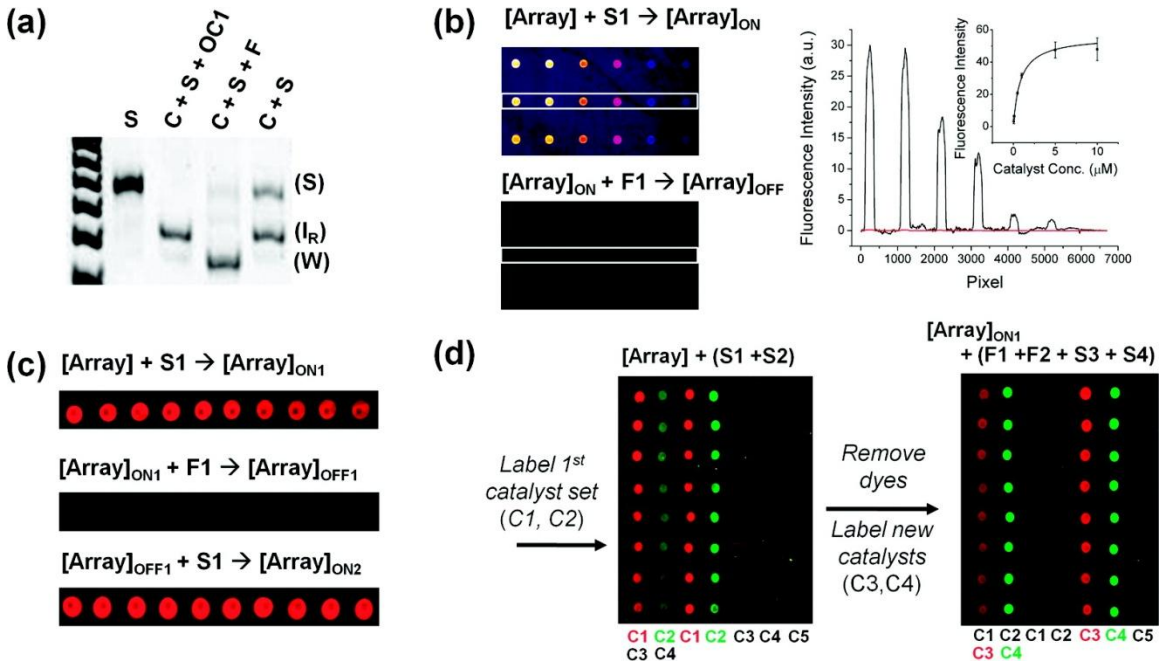

Figure 2.3 Multiplexed and reiterative fluorescence labeling of ssDNA catalyst. (a) Native PAGE-gel displaying DNA circuit reaction products: all components were reacted at a concentration of 200 nM. (b) A catalyst microarray possessing a gradient of spot concentrations (using 10, 5, 1, 0.5, 0.1, and 0.05 μM stock solutions) that was first labeled and then erased via the sequential addition of S (top image) and a fuel strand (bottom image). Images are rendered as heat maps. Plots of the averaged intensity profile of the boxed regions for the labeling (black line) and dye removal (red line) reactions are shown. Average spot intensities can be approximated by the Langmuir equation (inset, r^2 = 0.97). (c and d) Reiterative labeling of an individual (c) and multiple (d) arrayed catalysts. The reactions performed on the arrays are indicated above each image.

We next performed a series of imaging experiments that show DNA circuit complexes can be used to selectively label molecular markers on fixed and permeabilized HeLa cells. Background circuit reactivity was first tested by incubating cell samples with a quenched S complex (100 nM) for 1.5 h. The resulting images show no discernible fluorescence signals and possess signal to background ratios of 1 (Figure 2.4a), implying that the substrate complexes have exceptionally low background reactivity with cells. We

attribute this property to the high stability of the duplexed substrate complex, the enhanced sequence specificity of strand exchange over a classical hybridization mechanism [88], and the ability to place dye and quencher molecules in close proximity to one another within the substrate complex.

To label markers on cells, we chose to target the DNA circuit complexes to a transfected and expressed green fluorescent protein construct (GFP-Z_E) so that circuit labeling and dye removal efficiencies could be benchmarked directly against an internal standard (Figure 2.4b). The catalyst strand was coupled to the GFP using DNA-conjugated artificial-protein-based polymers (Z_R-ELS_6-ssCat) that we have previously developed for protein-DNA labeling [84]. These polymers associate with the GFP-Z_E via a heterodimeric leucine zipper complex (Z_E/Z_R: K_D ~10-15 M). Thus, after incubating GFP-Z_E transfected HeLa cells with Z_R-ELS_6-ssCat and then washing the samples to remove unbound polymer, GFP-Z_E transfected cells can be labeled by a reaction of a circuit substrate complex that carries a Cy5 dye and a quencher. As seen in Figure 2.4b, cells that were successfully transfected with GFP-Z_E reacted with S to produce fluorescent signals in the Cy5 channel of the microscope. While cells that were not transfected did not exhibit fluorescence, clear linear correlations are observed between GFP-Z_E and circuit labeling intensities, yielding a correlation coefficient r = 0.95.

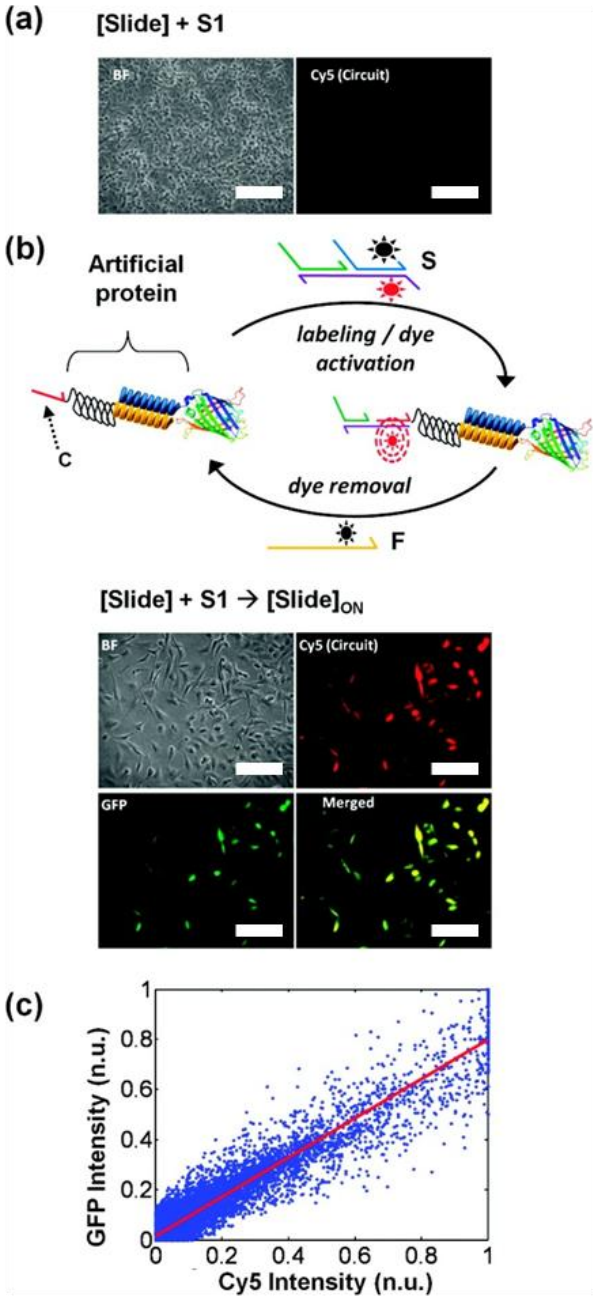

Figure 2.4 Selective targeting of an exogenously expressed GFP protein marker. (a) Bright field (BF) and fluorescence images of paraformaldehyde-fixed and permeablized HeLa cells reacted with a substratecomplex that incorporates a Cy5 and a quencher (Iowa Black). (b) Circuit-based labeling of GFP-Z_E transfected HeLa cells. The catalyst is attached to the GFP marker via a DNA-conjugated artificial protein (Z_R-ELS_6-ssCat) that is introduced postfixation. (c) Intensity correlation analyses of GFP-Z_E and DNA circuit (Cy5) signals. The correlation coefficient, r, for the fitted line is 0.95. (Scale bars: 100 μm)

We also tested whether molecular markers could be labeled multiple times on a single sample of cells without loss of fluorescence signal intensities (Figure 2.5). After a first round of circuit labeling and imaging, fuel strands were added to remove Cy5 dyes from a sample of GFP-Z_E transfected cells. As in our microarray experiments, dye removal reactions are found to be efficient and yield signal to background ratios of 1. In addition, the transfected cells could be labeled and imaged a second time by incubating the sample with a fresh solution of substrate. Brightfield imaging showed that a small portion of the cells detached from the slide surface during our manual washing and coverslip mounting procedures. Nevertheless, GFP and Cy5 signals remain highly correlated on a pixel-by-pixel basis after both rounds of fluorescence labeling ($r > 0.95$). Furthermore, strong correlations were found between cell images collected after the first and second dye labeling reactions are performed (i.e., between the ON1 and ON2 images in Figure 2.5 bottom). The Cy5 intensities of both images are linearly correlated and can be approximated by a line possessing a slope of 1.1. We therefore conclude that the Cy5 dyes can be coupled to the GFP-Z_E markers with near-identical efficiencies through sequential rounds of circuit-based labeling.

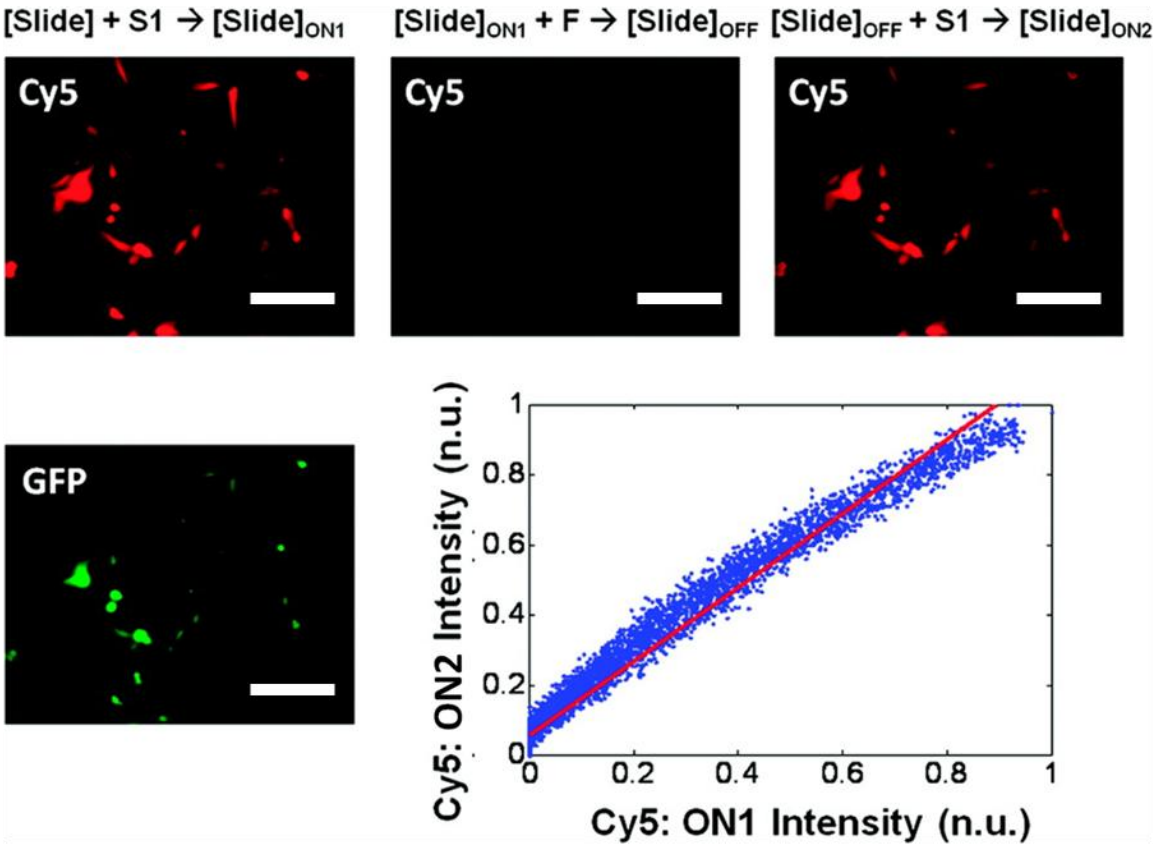

Figure 2.5 Reiterative circuit labeling of GFP-Z_E transfected cells. Sequential images of GFP-Z_E transfected cells where circuit complexes were used to label (ON1), erase (OFF), and relabel (ON2) the same sample of HeLa cells are shown. The pixel intensities in both Cy5 images (Cy5/ON1 and Cy5/ON2) are linearly correlated; fitted slope = 1.1, r = 0.92. (Scale bars: 100 μm)

2.4. Conclusions

In summary, we have demonstrated that DNA circuit complexes can be used as erasable molecular imaging probes. Here, the sequence-dependent specificity of the DNA circuit reactions facilitates multiplexed marker detection. The use of strand displacement mechanisms also allows fluorescence reporting complexes to be disassembled, and hence, new reporting complexes can be created and used to image additional sets of molecular

markers. Importantly, these reactions can be carried out at ambient temperature and in mild buffering conditions to minimize potential perturbations to a biological specimen. While such capabilities should offer opportunities to increase the number of molecular markers that one can examine on a single biological sample via fluorescence microscopy by at least a factor of 2 or 3, the next challenge will also be to develop diverse sets of molecular targeting agents (e.g., monovalent DNA-conjugated antibodies) that can facilitate efficient molecular marker recognition and react with the DNA circuitry reliably. Efforts to optimize syntheses of such agents are currently underway.

Chapter 3

Configuring Robust DNA Strand Displacement Reactions for *in situ* Molecular Analyses

The last chapter demonstrated the feasibility of activating and erasing fluorescent signals using engineered DNA complexes. Based on the results of that work, we hypothesized that simpler constructs may be capable of recapitulating this function of the substrate complexes in Chapter 2. In addition, smaller complexes may actually function better due to their ability to diffuse in and out of cellular environments more easily. New probe designs were therefore designed and evaluated by three main criteria: 1. Probes must activate fluorescence quickly and efficiently, 2. Erasing should also occur quickly with final signals < 10% of their activated intensity, and 3. Probes should be capable of multiplexed labeling and erasing.

To assess this, a two strand complex was created based upon the same criteria as the three-strand entropic circuit designs we previously utilized and tested them using the

transfectable GFP fusion system in Chapter 2. While the two-strand complex enhanced the activation reaction, it was unable to erase as efficiently as the three-strand complex. We hypothesized that the two-strand erasing had a strong reverse component, slowing the overall reaction. To overcome this, we devised a four-way branch migration scheme for two-strand complexes, which would create more "inert" products during erasing. This design was then able to activate and erase efficiently on fixed cell samples, permitting multiple activation and erasing reactions simultaneously to detect four unique targets.

This chapter has been adapted from:

Duose, Dzifa Y.; Schweller, Ryan M.; Zimak, Jan; Rogers, Arthur A.; Hittelman, Walter N.; Diehl, Michael R.; Optimized DNA Strand Exchange Reactions for In Situ Visualization of Biomarkers. *Nucleic Acids Res,*, 2012, 40 (7), 3289-3298.

3.1. Introduction

Recent advances in the field of DNA nanotechnology have facilitated the creation of various dynamic DNA and RNA complexes that can function as programmable logic gates [60, 65, 89], chemical amplifiers [71, 90], and reconfigurable molecular structures [91, 92]. A key feature of these complexes is that, instead of classical hybridization reactions, they can operate via a process called strand displacement – the exchange of oligonucleotides possessing partially or fully identical sequences among different thermodynamically-stable multi-strand complexes [93] (examples are shown in Figure 3.1 and Figure 3.2). Using this mechanism, long nucleic acid complexes possessing many matched base pairs can be hybridized and dehybridized multiple times at room

temperature. Moreover, since strand displacement reactions are sequence dependent and tend to be more sensitive to base mismatches than classical hybridization reactions [88], different dynamic complexes can be designed to operate independently of one another, or, alternatively, integrated into programmable reaction networks that can perform complex computations [66, 83, 94]. Such capabilities now offer opportunities to create new classes of molecular probe technologies for molecular-cell analyses.

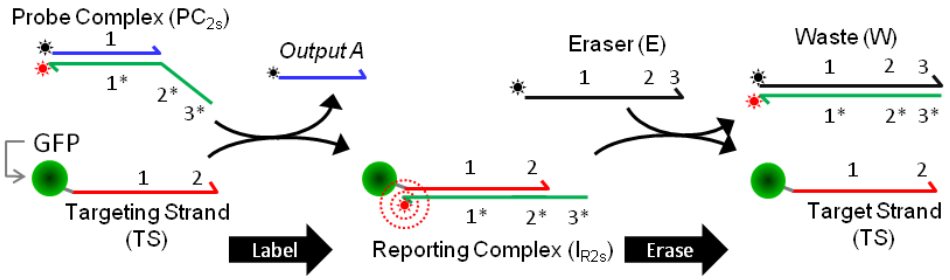

Figure 3.1 Labeling and erasing reactions for two-strand probe constructs.

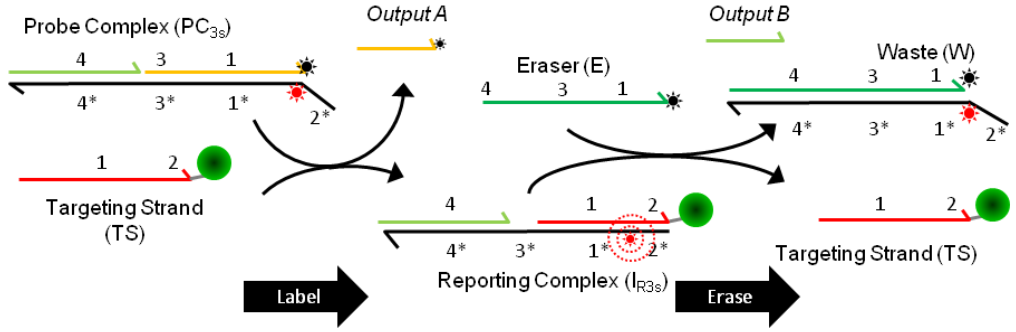

Figure 3.2 Labeling and erasing reactions for three-strand probe constructs.

The potential of dynamic nucleic acid complexes for various biological applications are beginning to be realized [72, 95, 96]. Engineered RNA hairpin devices

have been used as ‘smart’ therapeutic technologies that can selectively react with mutant RNA transcripts *in vitro*, and, in response, produce double stranded RNA polymers that trigger cell apoptosis [96]. Multiplexed (5-color) *in situ* detection of mRNA transcripts in fixed zebrafish embryos has also been demonstrated using similar hairpin systems [72]. Yet, despite these advances, the translation of dynamic oligonucleotide complexes towards such applications remains generically challenging. Most candidate probe constructions are first evaluated in a test tube where displacement reactions occur in homogeneously-mixed solutions [71, 85, 88]. The environment inside cells is much more complex and heterogeneous, and even if samples are fixed and permeabilized, issues surrounding the sample penetration and probe dispersion must now be addressed. Other environmental factors may also potentially interfere with the strand displacement process, and such effects could result in unwanted reverse or side reactions. For example, many dynamic DNA complexes are designed to consume a single-stranded input (a target) via toehold-mediated strand displacement. Here, a small single-stranded domain (~ 6 bp), called a toehold, is used to partially hybridize to the target and accelerate the initiation rates of the strand displacement reaction. Once the composite reaction is completed, a new, fully-duplexed complex is produced that is considered to be an ‘inert’ or unreactive byproduct since it no longer possesses a single stranded toehold domain. However, molecular crowding agents present in cells (e.g., proteins and other biomacromolecules) and enhanced concentration effects due to local confinement could potentially accelerate non-toehold-mediated strand exchange events between strands within duplexed DNA complexes and other unhybridized strands present in the sample [97]. In turn, the products formed from a displacement reaction would no long be fully inert. Thus,

addressing the unique challenges of detecting biomacromolecules within cells and tissues requires characterization of the strand exchange processes within these environments and the development of approaches to potentially circumvent these types of problems.

Our group is interested in developing dynamic DNA complexes that can function as reconfigurable (erasable) molecular imaging probes for *in situ* analyses of proteins [95]. These probes harness the strand displacement mechanism to selectively couple fluorophore-labeled oligonucleotides to DNA-tagged protein markers on cells. After these markers are visualized via fluorescence microscopy, a second strand displacement reaction is then performed to strip the dye-bearing strands from the reporting complexes that are coupled to the protein targets. With this capability, the total number of proteins that can be visualized on an individual sample is no longer limited by the spectral overlap of fluorescent dyes since the removal of dyes now allows new sets of protein markers to be detected on the same cell or tissue section. For this application, the strand displacement reactions driving the labeling and erasing of markers must be selective and efficient (high labeling and dye removal yields) to facilitate quantitation of marker levels and to ensure residual signals that remain on a sample after an erasing step do not compromise subsequent image analyses of other proteins. As with other probe systems, the use of small diffusible probe complexes is likely important to ensure their even probe dispersion throughout a sample. Finally, because individual samples will be inspected multiple times, the time required to complete the entire reiterative marker imaging procedure should be kept to a minimum. Thus, short marker labeling and erasing reaction times are highly desirable.

Herein, we evaluate the *in situ* marker labeling / erasing efficiencies and reaction kinetics of three different erasable dynamic DNA probes that contain different numbers of component strands and reactive domains. We also compare the labeling and erasing performance of probes that react with their target strands via 3-way and 4-way strand displacement mechanisms. These analyses indicate that efficient *in situ* erasing of DNA-tagged protein targets requires probe constructions that minimize the ability of product complexes that are produced during erasure to react with and relabel free ssDNA targets within cells via strand exchange processes that are not mediated by a toehold. In particular, the avoidance of these processes appears to be important to facilitate the rapid diffusion of fluorophore-bearing strands out of the cells and into their surrounding medium during erasing reactions. Overall, we find that 3-way strand displacement reactions where a single strand liberates two different strands from labeled targets support efficient erasing. Alternatively, fluorescent dyes can be removed efficiently using probe designs that exchange nucleotides via 4-way strand displacement processes since these reactions produce much more 'inert' dsDNA products. While facilitating the optimization of probe designs for our application, such information should aid the development and implementation of other dynamic DNA systems that are designed for analogous biological applications.

3.2. Materials and Methods

3.2.1. General

Oligonucleotides were purchased from Integrated DNA Technologies (IDT). The protein target, a recombinant green fluorescent protein, GFP-Z_E, was produced using

standard cloning and cell transfection procedures. The C-terminal leucine zipper (Z_E) is used as an affinity tag for DNA labeling. This zipper associates strongly ($K_D \sim 10^{-15}$ M) with a complementary basic zipper (Z_R) that is incorporated into a DNA-conjugated artificial protein (Z_R-ELS_6-ssTS) as a gene fusion [98]. These polymers were produced according to previously reported procedures [84]. The GFP-Z_E offers some advantages as a protein target for the present study since it can be outfitted with ssDNA stoichiometrically. The Z_R-ELS_6-ssDNA polymer is not depicted in Figure 3.1 and Figure 3.2 for simplicity.

3.2.2. Probe Design

Probe sequences - probe complex (PC), eraser (E) and targeting (TS) strands - were designed using similar methods to those described in references [95]. All sequences, excluding those adopted from Zhang *et al.* [68], were selected using a custom MATLAB script that generates random domains of specified lengths having pre-determined GC% range, while excluding previously generated domains or other prohibitive sequences (i.e. G quadruplexes), and avoiding secondary structures (e.g. hairpins). The generated domains are ranked according to their normalized two-state hybridization energies with existing probe strands using mFold [99]. The domains are then screened through the BLAST database (http://blast.ncbi.nlm.nih.gov/) to minimize probe sequence homology with the mRNA transcriptome. The final domain sequences are then selected manually from this list and concatenated with other domains to create full oligonucleotide sequences that will be incorporated into a probe complex. Other global criteria such as temperature, strand concentration, and salt concentration are specified prior to domain

design. A listing of the oligonucleotide sequences for all complexes can be found in Table 3.1.

3 Strand Probe System (PC_{3s})*	
Strand	**Sequence (5′-3′)**
TS	CAT TCA ATA CCC TAC GTC TCC ATT TTT TTT TT/AmMC6/
Output A (OA)	CCA CAT ACA TCA TAT TCC CTC ATT CAA TAC CCT ACG/IAbRQSp/
Output B (OB)	CTT TCC TAC ACC TAC GTC TCC AAC TAA CTT ACG G
Cy5LB-1	/Cy5/TGG AGA CGT AGG GTA TTG AAT GAG GGC CGT AAG TTA GTT GGA GAC GTA GG
Eraser (E)	CCT ACG TCT CCA ACT AAC TTA CGG CCC TCA TTC AAT ACC CTA CG/IAbRQSp/
2 Strand Probe System (PC_{2s} 3-way)*	
TS	/AmMC6/TTT TTT TTT TCG AGA TGC CTT ACA GTA GGT TGG A
Output A (OA)	CGA GAT GCC TTA CAG TAG/IAbRQSp/
Cy5_Dye	/Cy5/ACG ATG TCC AAC CTA CTG TAA GGC ATC TCG
Eraser (E)	CGA GAT GCC TTA CAG TAG GTT GGA CAT CGT/IAbRQSp/
2 Strand Probe System (4-way)*	
TS	/AmMC6/TTT TTT TTT TCG AGA TGC CTT ACA GTA GGT TGG A
Output A (OA)	/IAbRQ/AGA TGC CTT ACA GTA G
Cy5_Dye	ACG AAC TCC AAC CTA CTG TAA GGC ATC T/Cy5Sp/
Eraser (E)	/IAbRQ/AGA TGC CTT ACA GTA GGT TGG AGT TCG T
EraserP (EP)	ACT CCA ACC TAC TGT AAG GCA TCT

Table 3.1 List of oligonucleotide sequences used in design of DNA-circuits. /AmMC6/ represents an amino modifier; /Cy5/ or /Cy5Sp/ indicate a Cy5 fluorophore; /Cy3/ or /Cy3Sp/ indicate a Cy3 fluorophore; and /IAbFQ/ or /IAbRQ/ represent an Iowa Black Quencher for the green to pink or red spectral ranges, respectively

Strand	Sequence
2 Strand Probe System2 (PC_{2s} 4-way)	
TS	/AmMC6/TTT TTT TTT TGT GTA CCG GAA ACA TCG GCG AAT TAG
Output A (OA)	/IAbFQ/GTG TAC CGG AAA CAT CGG
Cy3_Dye	CTT GTC AAT TCG CCG ATG TTT CCG GTA CAC/Cy3Sp/
Eraser (E)	/IAbFQ/GTG TAC CGG AAA CAT CGG CGA ATT GAC AAG
EraserP (EP)	CTA ATT CGC CGA TGT TTC CGG TAC AC
2 Strand Probe System3 (4-way)	
TS	/AmMC6/TTT TTT TTT TGG CCA CCG AGA CAA TAC GCA GGA CCC
Output A (OA)	/IAbFQ/GGC CAC CGA GAC AAT ACG
Cy5_Dye	CCT TAA GTC CTG CGT ATT GTC TCG GTG GCC/Cy5Sp/
Eraser (E)	/IAbRQ/GGC CAC CGA GAC AAT ACG CAG GAC TTA AGG
EraserP (EP)	GGG TCC TGC GTA TTG TCT CGG TGG CC
2 Strand Probe System4 (4-way)	
TS	/AmMC6/TTT TTT TTT TGA TAT CAA GCT GCT CTG GGT ATG C
Output A (OA)	/IAbFQ/GAT ATC AAG CTG CTC T
Cy5_Dye	AAT CCT ATA CCC AGA GCA GCT TGA TAT C/Cy3Sp/
Eraser (E)	/IAbFQ/GAT ATC AAG CTG CTC TGG GTA TAG GAT T
EraserP (EP)	GCA TAC CCA GAG CAG CTT GAT ATC

Table 3.1 Cont.

Fluorophores (Cy3 or Cy5) and quencher molecules (Iowa Black FQ or RQ) were incorporated in opposing strands at positions that minimized their intermolecular distances in both the probe complexes and their waste products (Figure 3.1 and Figure 3.2). For the 3-strand probes (PC_{3s}), this requires that the dye molecule is positioned internally within the longer strand of the complex (Figure 3.2).

3.2.3. Cell Labeling and Erasing Procedures

CHO cells were grown on glass coverslips in F12 media supplemented with 10% (v/v) FBS. After 24 hours, the culture medium was replaced and the cells were transiently transfected with vector containing the GFP-Z_E construct using Fugene (Roche) according to the manufacturer's protocol. Cells were cultured for an additional 12 hrs to allow for GFP-Z_E production. The cells were then fixed using freshly prepared 4% (w/v) paraformaldehyde for 30 min. Activated aldehydes resulting from the fixation procedure were quenched using 1 mg/ml $NaBH_4$ for 5 min at room temperature. Afterwards, the cells were permeabilized using 0.2% (v/v) Triton X-100, washed twice with PBS, and stored overnight a 4 C.

Prior to cell labeling experiments, the coverslips were rinsed twice in PBS, dried under an airstream, and then affixed to custom-fabricated micro-well chambers (10-round wells with 0.36 cm^2 culture area and culture volume of 400 µl) using a precision-cut double sided adhesive film. The cells were re-hydrated with PBS prior to the labeling procedure. To minimize non-specific binding of the Z_R-ELS_6-ssTS, cells were first blocked for 2 hrs using a solution containing 1% (w/v) BSA, 1 mg/ml Herring Sperm DNA and 0.5 µM polyT DNA in PBS. The cells were then incubated with a 400 nM solution of Z_R-ELS_6-ssT for 2 hours, and washed twice with PBS.

Cell labeling experiments were performed by incubating cells with solutions of 100 nM probe complexes in TAE buffer supplemented with 12.5 mM Mg^{+2}. Probe deactivation / erasing reactions were performed using 1 µM of the eraser strands (E_s) or complexes (E_c). All cell labeling / erasing reactions were carried out for 2 hours at 30º C using a rotating incubator shaker (200 rpm), except for the kinetic experiments where the

reactions were performed directly on the microscope at room temperature and without shaking.

Cells were imaged using an inverted Nikon microscope outfitted with a 40x 0.95 N.A. objective, electronic shutters, and a 14-bit depth EMCCD camera (LucaR; Andor South Windsor, CT). A mechanical transitions stage and electronic focusing mechanism was used to collect 5-10 different image fields for each sample. Images were processed using Nikon (NIS image) or ImageJ software, and are presented as heat maps since this rendering enhances the contrast of low-level, remnant fluorescence signals within the 'erased' images. Average cell intensities were determined using an algorithm in NIS image that marks the boundary of a selected cell and calculates the average pixel intensity within that region.

3.3. Results and Discussion

One of the goals of this study is to define simple probe designs containing a minimal number of component strands that can function effectively as erasable molecular imaging probes. Since this capability allows fluorophores to be removed from a sample and then the same color dyes to be reused to label new proteins, this functionality should facilitate more comprehensive biomarker analyses where multiple sets of proteins are visualized on the same cell or tissue samples. Due to the spectral overlap of fluorescent dye molecules, the number of markers that can be visualized using fluorescence microscopy and permanently stained samples is typically limited to three to five proteins. However, this number can be at least doubled using our reiterative labeling and erasing procedure, and hence, one could detect at least 6-10 proteins using standard, 3-5 color

fluorescence microscopy procedures, while avoiding the use of temperature, bleaching, caustic chemicals to remove dyes from sample. Moreover, several spectral deconvolution microscopy approaches have been developed that offer opportunities to increase the number of markers that are detected during a single round of fluorescence microscopy. Combining this capability with our reiterative labeling techniques could therefore facilitate even more extensive *in situ* analyses of protein markers.

Our previous report showed that a class of 3-strand DNA complexes (PC_{3s}) that have been previously integrated into catalytic networks [68] can also function effectively as erasable molecular imaging probes, providing linearly-correlative labeling intensities and efficient > 95% dye removal [95]. Despite this success, we believe our application would benefit from the development of somewhat smaller, 2-component probe complexes that incorporate terminal (3′ or 5′) dye molecules instead of the internal dyes in our prior design. There was concern that the internal dye placement could potentially interfere with the strand-displacement process and restrict the types of dyes that can be incorporated into a probe. To address this issue, we created several DNA probe complexes composed of two partially complementary DNA oligonucleotides (PC_{2s}; Figure 3.1). The marker labeling and erasing reactions for the original 3-strand complexes are illustrated in Figure 3.2. In both cases, the probes react with their ssDNA targets (TS) via toehold-mediated strand displacement to produce a fluorescent reporting complex (I_R) containing an unquenched fluorophore. As a result, molecular targets that are conjugated with TS can be visualized using fluorescence microscopy. Analogously, the fluorescent reporting complexes (I_{R2s} and I_{R3s} for the 2- and 3-strand probes respectively) can be displaced using a single-stranded eraser oligonucleotide (E). In

principle, the quenched waste product (W) of this reaction can then be washed off the sample. Of note, the strand displacement in the labeling and erasing reactions of each probe system proceed via a 3-way branch migration process [100], where ssDNA components displace one another.

The present PC_{2s} probes contain three distinct domains (Figure 3.1): an 18 bp domain that is completely hybridized (domains 1 / 1*) and two 6 bp toehold domains that are positioned adjacent to one another at one end of the complex (domains 2* and 3*). Reporting complexes of the PC_{2s} system are formed through a toehold-mediated exchange reaction that is initiated by binding of the probe complex to domain 2 of the TS strand. However, unlike the 3-strand complexes whose reaction is completed by the release of an output strand from a second toehold domain (output A releases from domain 3 in Figure 3.2), the PC_{2s} probes were designed so that TS displaces the output strand (output A) within PC_{2s} completely during the marker labeling reaction. Similarly, the I_{R2s} reporters are disassembled / erased through a displacement reaction where an eraser strand (E) binds to toehold 3* on I_{R2s} and then displaces the TS completely from the reporting complex. Since neither reaction require the dehybidization of a toehold domain, the forward reactions during marker labeling and erasing both result in the accumulation of 6 additional matched bps. This design was originally chosen so that each reaction would be energetically favorable. Additionally, since the release of outputs from toehold domains could potentially lower probe-target strand exchange rates, we generally expected that this feature would accelerate our labeling and erasing reactions.

Another important distinction between the PC_{2S} and PC_{3s} probes is that their erasing reactions produce different numbers of products. For the 3-strand probes, the

marker erasing involves a reaction where E binds to a 4 bp domain in I_{R3s} and then displaces two strands (output B and TS) from the reporting complex. Thus, even though this reaction results in a net loss of 2 matched bps in the system, there is an increase in configurational entropy as the reaction proceeds forward [68]. An important consequence of this design for our application is that reverse (relabeling) reactions can only occur if both B and TS bind to W simultaneously.

Finally, in an attempt to drive each reaction to complexion, our protein labeling and erasing reactions all use relatively high concentrations of E (1 μM) relative to TS (~2 nM in the total reaction volume). Thus, so long as the reactions of probes with their targets mimic the behavior one would expect in a homogenously-mixed solution at these concentrations, each probe reaction should be able to reach an equilibrium distribution where the vast majority of the TS strands (>>95%) were either incorporated into an I_R complex, or erased. In principle, such behavior should support efficient labeling and erasing.

To evaluate the labeling and erasing performance of the PC_{2s} and PC_{3s} probes, we performed two sets of *in situ* cell imaging experiments where expressed GFP proteins within fixed and permeabilized CHO cells were first outfitted with TS strands using an ssDNA-artificial protein conjugate and then reacted with either 2- or 3- strand probe complexes that incorporate Cy5 fluorophores (Figure 3.3 and Figure 3.4). In each case, labeling intensities were evaluated after the probes were allowed to react for a period of 2 hours. For practical purposes, reactions will likely need to be completed within a short time frame to implement our reiterative marker imaging technique.

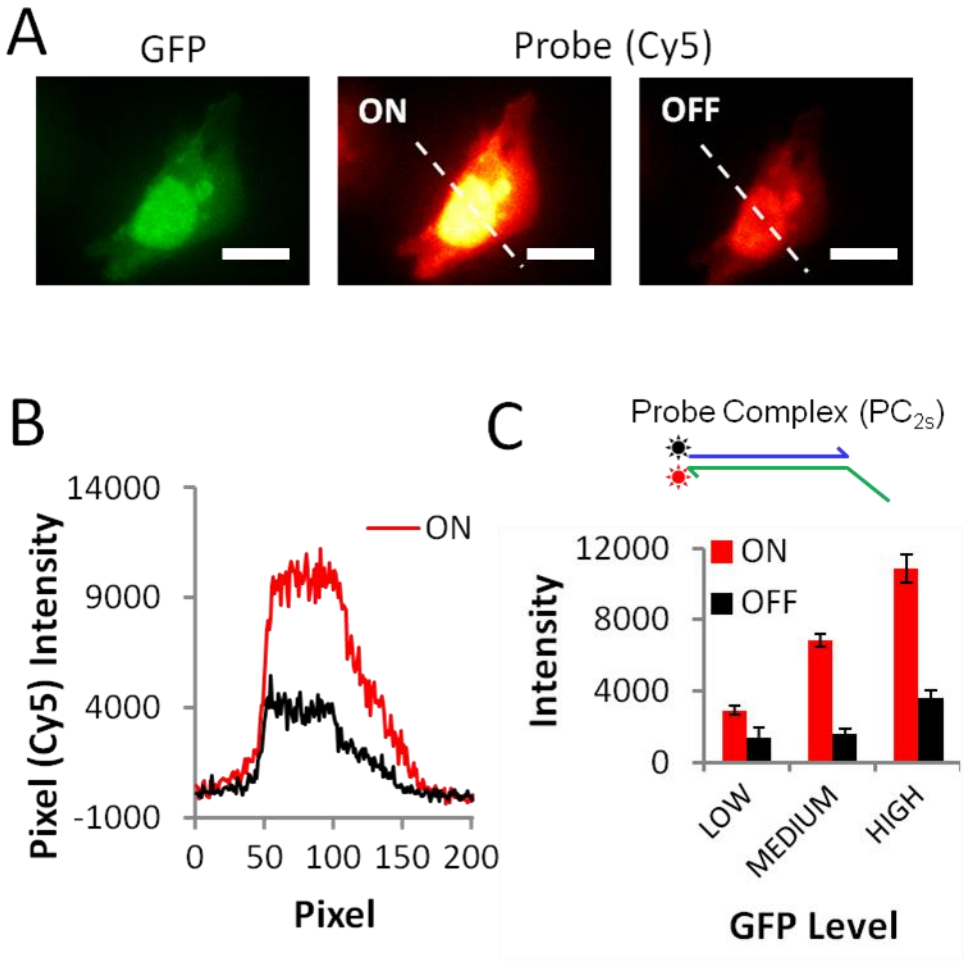

Figure 3.3 Labeling and removal of Cy5 fluorophores from protein markers via strand displacement reactions of a two-strand probe complex (PC_{2s}). (A) Selective labeling of expressed GFP proteins in CHO cells. The images display a strong correspondence between the GFP and two-strand probe (Cy5) signals; pixel intensities of the GFP and Cy5 signals are linearly correlated (r^2=0.94). However, OFF signal intensities indicate ~20% of active Cy5 dye remains on the cells after the erasing reaction. (B) Pixel intensities for cross section indicated in the probe images in A for both the ON and OFF states of the cells. (C) Histogram of the average Cy5 signal intensities for the ON and OFF states of 20 cells. Cells are grouped based on their GFP whole-cell fluorescence intensities (low, medium and high levels correspond to 2000–5500, 5500–10000 and 10000–15800 intensity units, respectively). (Scale bars: 20 μm)

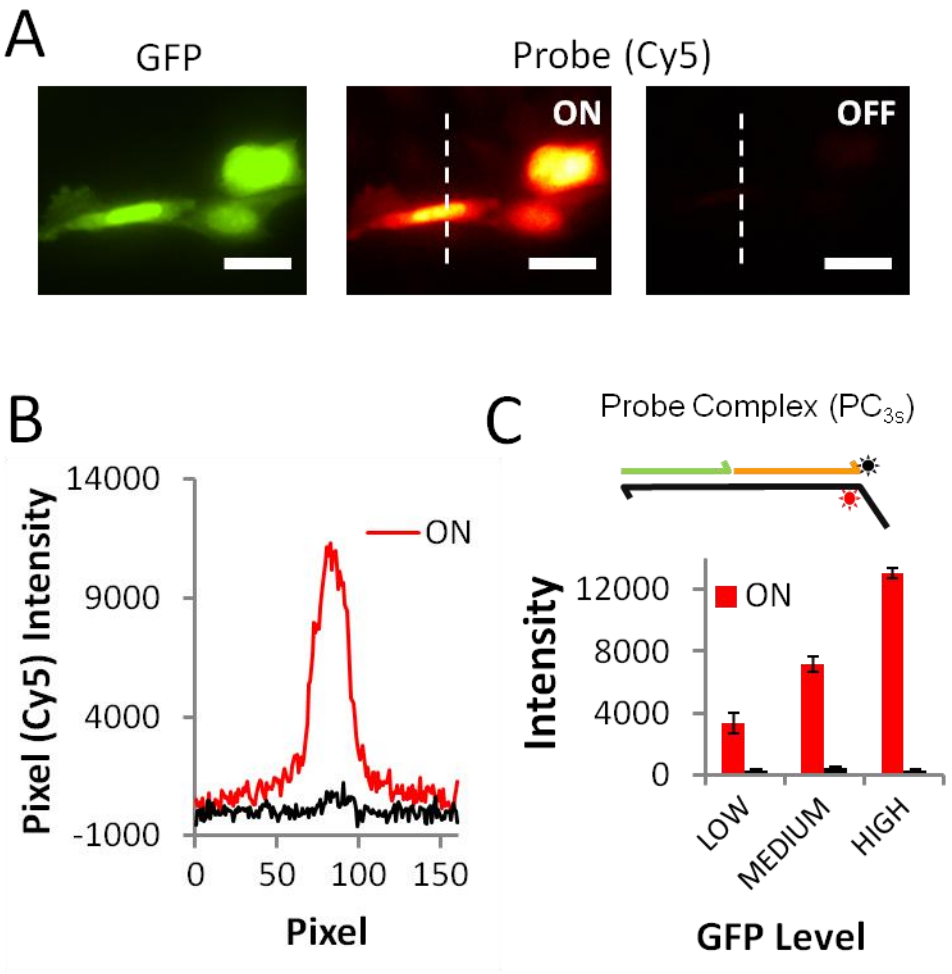

Figure 3.4 Labeling and removal of Cy5 fluorophores from protein markers via strand displacement reactions of a three-strand probe complex (PC_{3s}). (A) Selective labeling of expressed GFP proteins in CHO cells. The OFF reactions are now efficient, and yield a signal (cell intensity) to background (slide surface intensity) ratio of 1.08. The ratio of labeled/erased probe intensities, or ON/OFF ratio, is 28.6. This result is reflected in (B) pixel intensities for cross sections as well as (C) histogram of average, whole-cell Cy5 signal intensities for 20 cells in their ON and OFF states. (Scale bars: 20 μm)

Comparisons of GFP and Cy5 signals produced after a labeling reaction show that both types of probes can selectively couple fluorophores to DNA-conjugated proteins on cells. In each case, intracellular distributions of the GFP molecules (e.g., nuclear -

cytoplasm signal ratios) are reproduced, while the cells that were not transfected in the sample are not labeled. Yet, despite their similar protein labeling performance, the erasing reaction of the PC_{2s} probe system (Figure 3.3) is found to be much less efficient than the PC_{3s} system (Figure 3.4). Here, the erasing reaction that disassembles the 2-strand I_{R2s} reporter results in residual Cy5 signals that range from 20 to 30% of the signal amplitudes produced by the prior marker labeling reaction, yielding an average ON/OFF ratio of 3.14. Furthermore, these unwanted 'OFF-state' signals are positively correlated with 'ON-state' Cy5 intensities (Figure 3.3C, *inset*). The erasing performance of the 3-strand complex is much better (Figure 3.4). As in our prior work [95], 'OFF-state' signals with the 3-strand complexes can barely be detected over background autofluorescence of the unlabeled cells ('OFF'-signals are only 1-3 times the RMS noise of the slide background), and average ON/OFF ratios are much higher (26.8). Thus, of the two different probes systems evaluated above, only the 3-strand systems can function effectively as an erasable molecular imaging probe, suggesting that for 2-strand complexes, 3-way displacement reactions are insufficient for dye removal and erasing.

We next characterized the *in situ* marker labeling and erasing kinetics of the PC_{2s} and PC_{3s} systems by monitoring the rates that probe (Cy5) signals colocalized to, and then are removed from, expressed GFP molecules that are labeled with single TS strands via a DNA-conjugated protein polymer (Figure 3.5). For these analyses, cells were imaged every 5 minutes during the marker labeling / erasing procedures. Analyses of probe intensities show that TS-tagged GFP proteins are labeled rapidly by the 2-strand probes, and that probe signals saturate within 20 min, even for the cells possessing the highest GFP expression levels. The PC_{3s} labeling reactions were somewhat slower than

the PC_{2s} reactions, indicating that the release of the output strand from domain 3* in I_{R3s} and the accumulation of 2 instead of 6 base pairs during this forward reaction affects the rates that markers are labeled. Nevertheless, Cy5 intensities reached their plateau levels in less than 1 hour, except for the very brightest cells within the sample, and are linearly correlated with the GFP intensities on a pixel-by-pixel basis, as was the case in our prior report [95]. Thus, so long as the PC_{3s} probes are allowed to react for a sufficient period of time, both probe complexes can support efficient (relatively fast and proportionally correlative) marker labeling. Yet, Figure 3.5 also shows there are significant differences in the erasing kinetics of these systems. Although Cy5 fluorescence intensities decrease rapidly after the initial addition of eraser strands (E) in both cases, the PC_{2s} system erases more slowly than the PC_{3s} system. In addition, the PC_{2s} erasing reaction slows appreciably after a period of ~20 minutes, and significant Cy5 signals remain on the sample after the full 2-hour incubation period. It is our experience that > 12 hour incubations are necessary to achieve significantly lower fluorescent levels. In contrast, fluorescence intensities drop rapidly to a value that is less than 5-10% of their ON-state values within minutes during the reaction of E with 3-strand I_{R3s} complex.

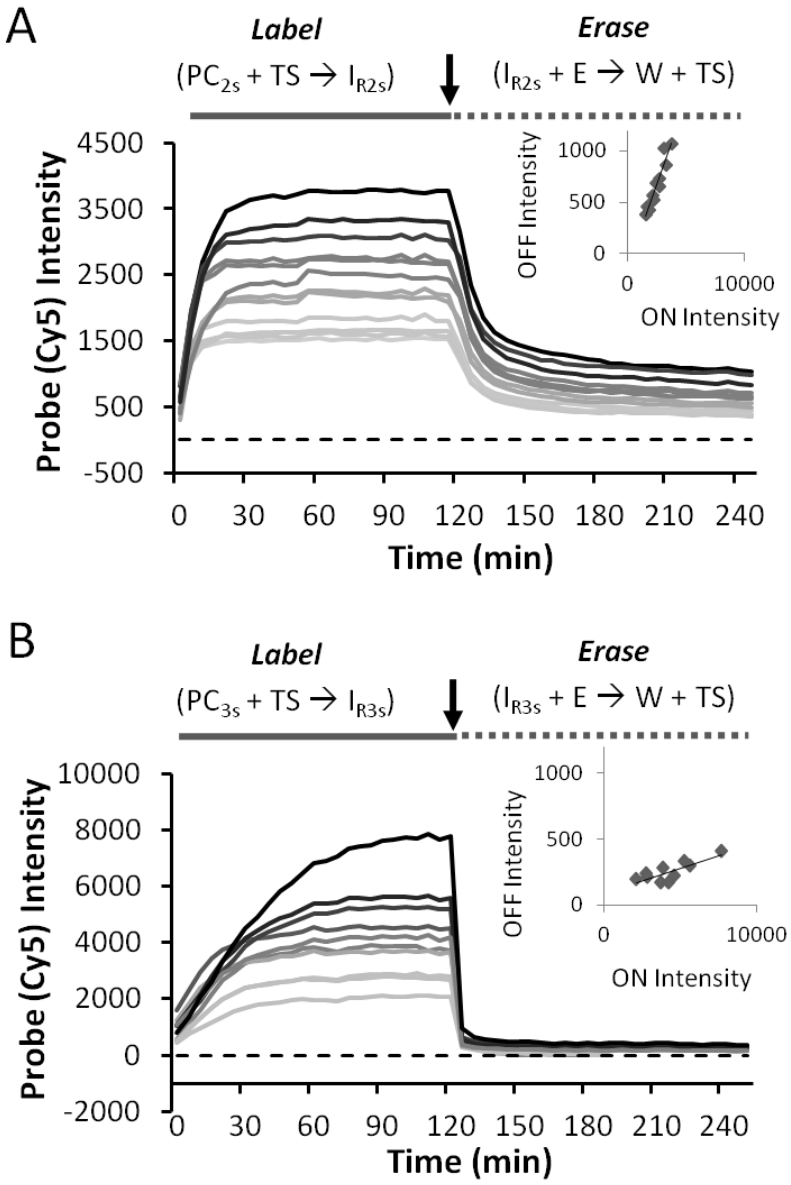

Figure 3.5 Kinetics of DNA strand displacement reactions on fixed cells. (A) Labeling and erasing reactions of a two-strand probe complex (PC_{2s}). Each curve represent the average intensity of an individual cell within the sample. The erasing reactions are inefficient and significant signals remain on the sample even after a 2-h incubation period. (B) Labeling and erasing reactions of a three-strand probe complex (PC_{3s}) showing rapid and efficient erasing. The arrows in each plot indicate the time point where the labeling reactions were stopped and the erasing reactions were initiated.

To gain further insight into the difference between the erasing behaviors of the PC_{2s} and PC_{3s} probes, we repeated the *in situ* kinetics experiments using eraser strands that do not contain quencher molecules (Figure 3.6). With this modification, duplexed I_R, W complexes, and other unidentified intermediates should all produce fluorescence signals if they remain bound to their targets or trapped within the cells during the erasing reaction. As shown in Figure 3.6B, the kinetic curves measured for unquenched 3-strand erasing reaction (I_{R3s} + E → W + output 2) are very similar to the original plots in Figure 3.5B. We therefore conclude that W is able to release from the 2* toehold domain within TS during the erasing reaction. The lack of persistent fluorescence in this setting rules out the possibility that the dye molecule remains tethered to the GFP target but is quenched by an interaction with E. The rapid drop of Cy5 signals during the erasing reaction in the 3-strand reporter complex setting suggests that the released W complex can freely diffuse out of fixed cells. In contrast, removing the quencher from reactions of 2-strand reporter complexes I_{R2s} with E results in appreciable remnant Cy5 signals on GFP transfected cells due to incomplete erasing (Figure 3.6A). In fact, the residual Cy5 signal remaining on the sample during the erasing reaction ("OFF-state") is nearly 75% of that seen for in the "ON-state".

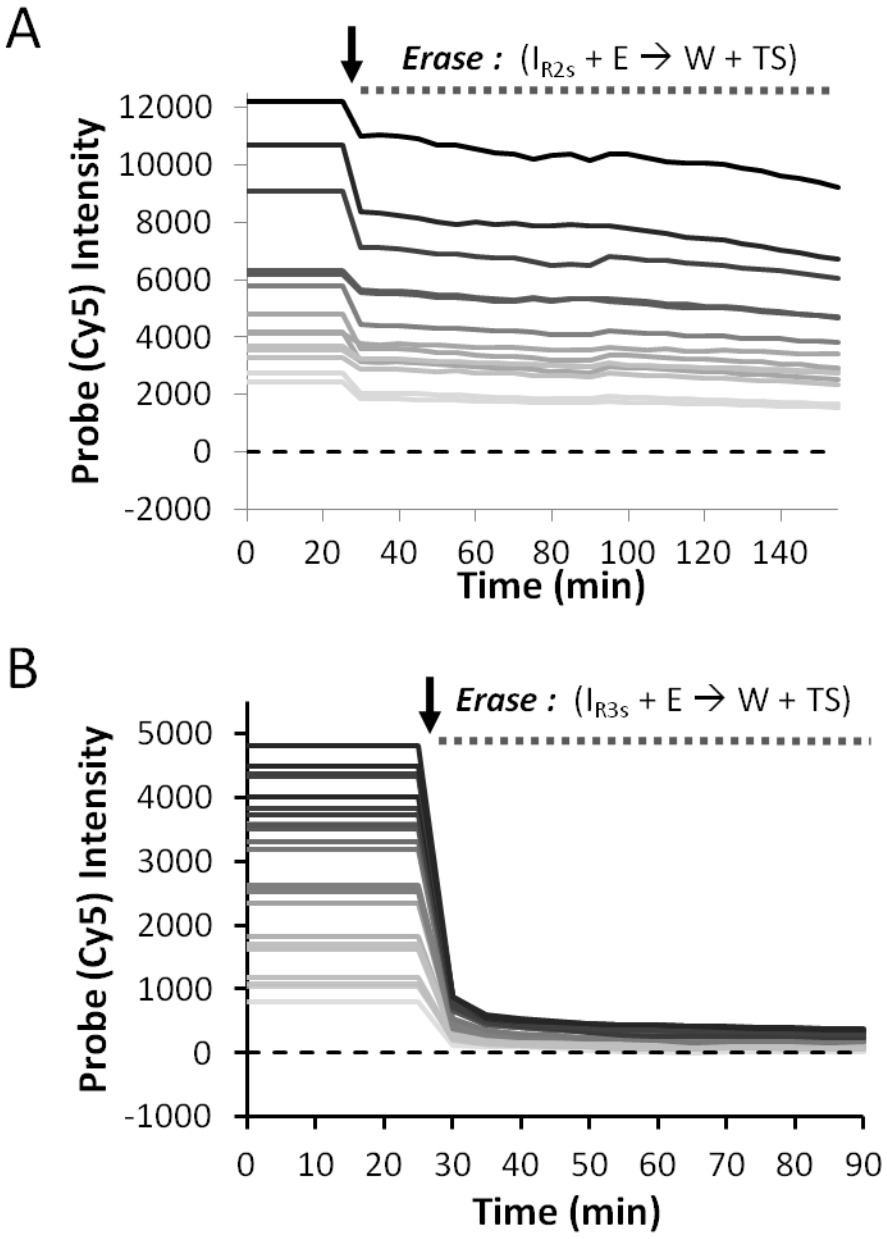

Figure 3.6 Erasing kinetics using eraser strands (E) that do not incorporate quencher molecules. (A) Erasing reactions of a two-strand complex (PC_{2s}) showing appreciable (~75%) residual signals. (B) Erasing reactions of a three-strand complex PC_{3s}.

Experiments where the quencher is removed from E suggests that the W complexes are kinetically trapped within the cells' volume and do not diffuse beyond the cell boundary within this time period. This behavior is somewhat surprising given the high E concentrations (1 μM) and reaction volume (100 μL) used for the erasing experiments. With such an excess of E relative to the total number of I_{R2s} reporting complexes on the cells, there should be a strong driving force to push the erasing reaction forward. However, given the compartmentalization of the GFP targets within the cells, the local concentration of TS and W can be quite high: 1,400 – 15,000 GFP/μm^3 or 25 – 250 μM according to GFP intensity analyses. Thus, high local target concentrations could serve to drive the reverse (relabeling) reaction (W + TS → I_{R2s} + E). Of note, the analogous reverse reaction (I_{R2s} + output A → TS + PC_{2s}) could be affecting the marker labeling step. Yet, the probe labeling intensities appear to saturate rapidly (< 30 min) and still yield marker intensities that are linearly correlated with GFP levels.

The *in situ* kinetic analyses of the PC_{2sC} erasing reactions with and without quencher in Figure 3.5A and Figure 3.6A indicate that, while significant numbers of I_{R2s} reporters remain on the cells after 2 hour incubations, fluorescently active W complexes are also somehow trapped within the cells. Because TS does not release from the waste complex via toehold dehybridization in this reaction (TS is displaced completely from the complex), we attribute that majority of the signal intensities present in Figure 3.5A to fluorescently-active W complexes as opposed to an unidentified, metastable intermediate-state complex. Importantly, we also find that TS-labeled GFP proteins can be selectively labeled using *in situ* reactions where the cells are simply incubated with a quenched W complex (Figure 3.7). This would suggest that the W complex can directly bind to the TS

strand and become fluorescently activated through a non-toehold mediated exchange of oligonucleotides. Thus, it appears that in the 2-strand reporter setting (but not in the 3-strand reporter setting), there may be some reverse (relabeling) reactions occurring during the erasing reaction resulting in a decreased diffusion rate of the W complex from the cell.

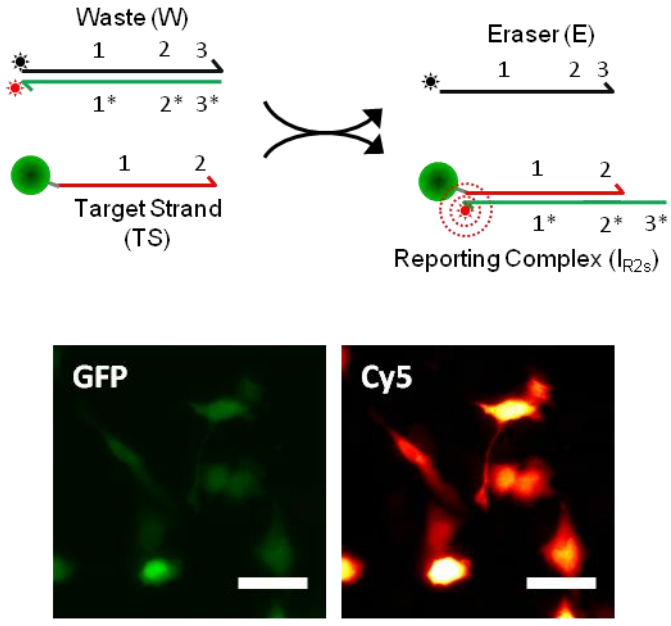

Figure 3.7 Labeling of DNA-tagged GFP proteins through non-toehold mediated exchange via a reaction of TS with a quenched W complex as a probe. The concentration of the W complex was 200 nM, and the reaction was performed without adding E to the reaction mixture. 4-way branch migration reactions facilitate efficient signal erasing. (Scale bars: 50 μm)

An important consequence of non-toehold-mediated exchange reactions is that they can influence the effective, diffusive mobilities of oligonucleotides and DNA complexes within cell samples. Through this reaction, strands containing the fluorophores will interact transiently with multiple immobilized TS strands as they diffuse towards the

cell boundary. These interactions can therefore lower the rates at which the fluorophore containing strands are released from the sample. While this effect limits removal of dyes from cell samples during our marker erasing procedure, analogous reverse and side reactions could potentially influence abilities to integrate the reactions among different dynamic DNA complexes for other biological detection applications. Nevertheless, the analyses of the 3-strand probe complexes (PC_{3s}) indicate these issues can be mitigated using probe designs where the reaction of a probe with its DNA target is entropically-favorable and two output strands are produced by the reaction. In the case of the 3-strand probe complexes, both outputs must bind to the W complex simultaneously in order to produce a fluorescently active complex that is stably bound to TS. As a result, W complexes are less likely to reassociate with TS and can more readily diffuse out of the cells. Alternatively, efficient marker erasing can be achieved using probes that react via 4-way branch migration processes. This mechanism produces two fully-duplexed waste products instead of only one, and hence, both products are much less likely to exchange their strands, but also, should be less reactive towards other oligonucleotide complexes present within cells. While such control has allowed our group to create a series of erasable imaging probes that facilitate the reiterative labeling of cell samples, overall, we anticipate these adaptations will be generically important to the development of other probe technologies that harness the unique functionalities of dynamic DNA complexes.

With the hypothesis that non-toehold-mediated exchange of oligonucleotides between W and TS complexes reduces the erasing performance of the PC_{2s} probes systems, we next tested whether a fully-erasable probe composed of only two different strands could be created that exchange nucleotides via 4-way [101] as opposed to 3-way

strand displacement mechanisms. To do so, the original 2-strand complexes were modified such that they still bind the same TS sequence, but leave a second 2 bp toehold unhybridized within the I_R complex (domain 4 in Figure 3.8A). As a result, while the labeling reaction proceeds near-identically to that of the original PC_{2s} systems, the 4-way branch migration processes of the erasing reaction now produces two reaction products that are fully duplexed (i.e., TS is incorporated into a duplexed complex after the reaction). In this case, probe erasing rates and efficiencies are quite similar to those of the 3-strand complexes. Measured 'OFF-state' intensities and reaction rates are now very close to those of the PC_{3s} system (Figure 3.8B and Figure 3.8C). Thus, 4-way branch migration reactions can be used as an alternative to the multi-strand release in the PC_{3s} systems in order to create more efficient erasable probes.

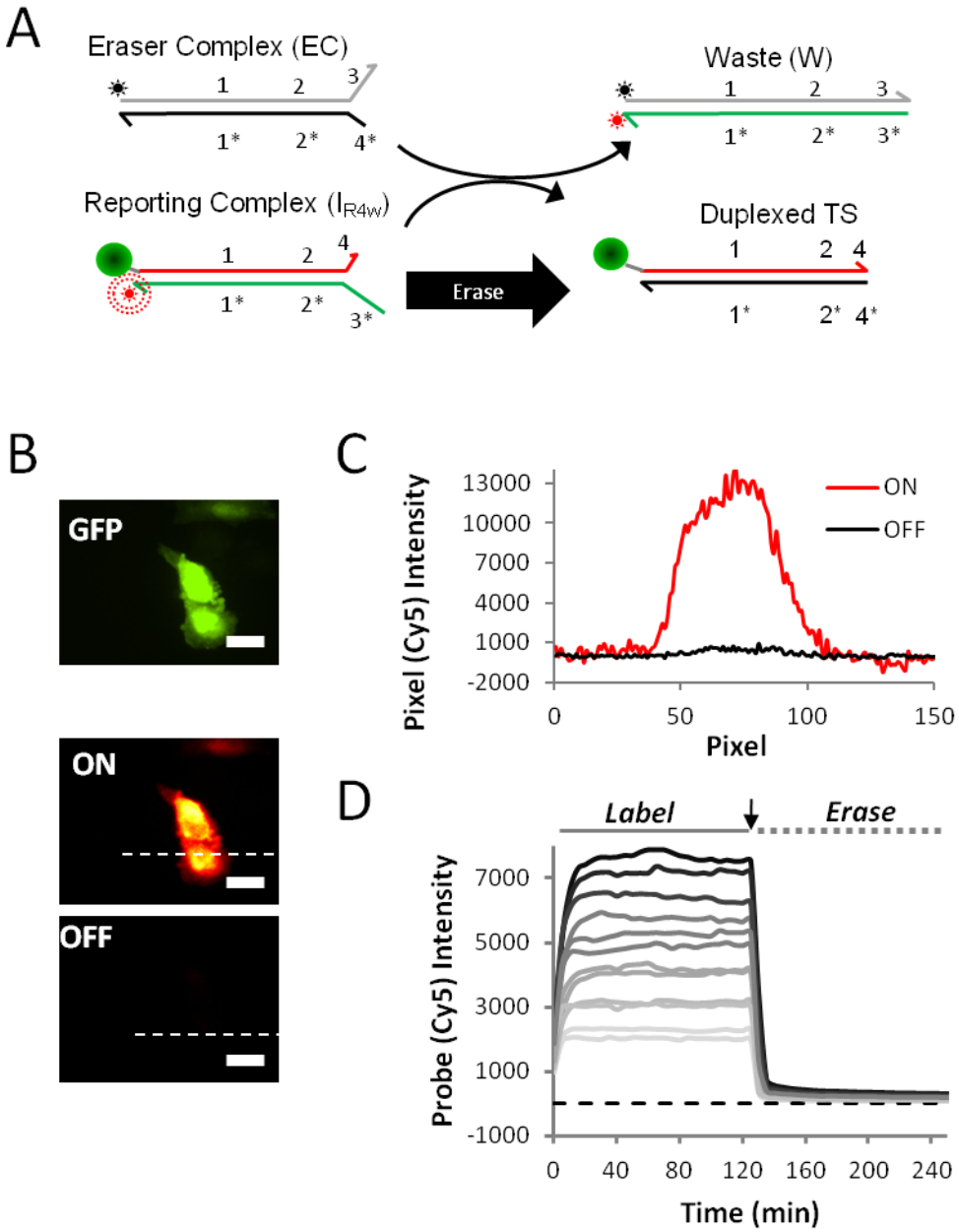

Figure 3.8 Labeling and removal of Cy5 fluorophores from protein markers using the two-strand probe complex that exchanges via a four-way branch migration process. (A) A scheme depicting the modified erasing reaction. (B) Selective labeling and erasing of expressed GFP proteins in CHO cells. (C) Pixel intensities for cross sections indicated in (B). (D) Kinetics of DNA strand displacement reactions on fixed cells showing that four-way branch migrations facilitate efficient marker erasing. (Scale bars: 20 μm)

To further evaluate the utility of the 4-way branch migrations for marker erasing, we also labeled expressed GFP-Z_E with mixture of four different DNA conjugates in a single incubation step at approximately equal molar concentrations (Figure 3.9). The cells were then reacted simultaneously with two probes that couple Cy3 and Cy5 dye molecules to the GFP targets bearing the TS1 and TS2 strands. In a single incubation step, these two dyes can be removed from the sample using the 4-way displacement mechanism, yielding average ON/OFF ratios that are > 10:1. Subsequently, two new complexes can be used to label GFP molecules coupled to the TS3 and TS4 targets in a second round of marker labeling. After this step, the probe Cy3 and Cy5 signals reappear at levels that are similar to those produced in the first labeling step. The second labeling produced Cy3 and Cy5 signals were 60% and 120% of their corresponding ON-state values (note: the dyes are now coupled to different TS strands on the sample; some of this variability can come from differences in the concentrations of the conjugated polymers used to couple the 4 different TS strands to the GFP target). Importantly, 'OFF'-state signals of the cells only constitute 2 – 6 % of the measured 'ON'-state signals produced by the second labeling reaction, and, hence, contribute little to the signals generated after the second set of TS markers were labeled. These experiments therefore indicate the 4-way displacement mechanism can be used to create multiple erasable probes that can be used to label multiple sets of markers with the same color dyes, or even to perform replicate measurements using different probes for standardization purposes.

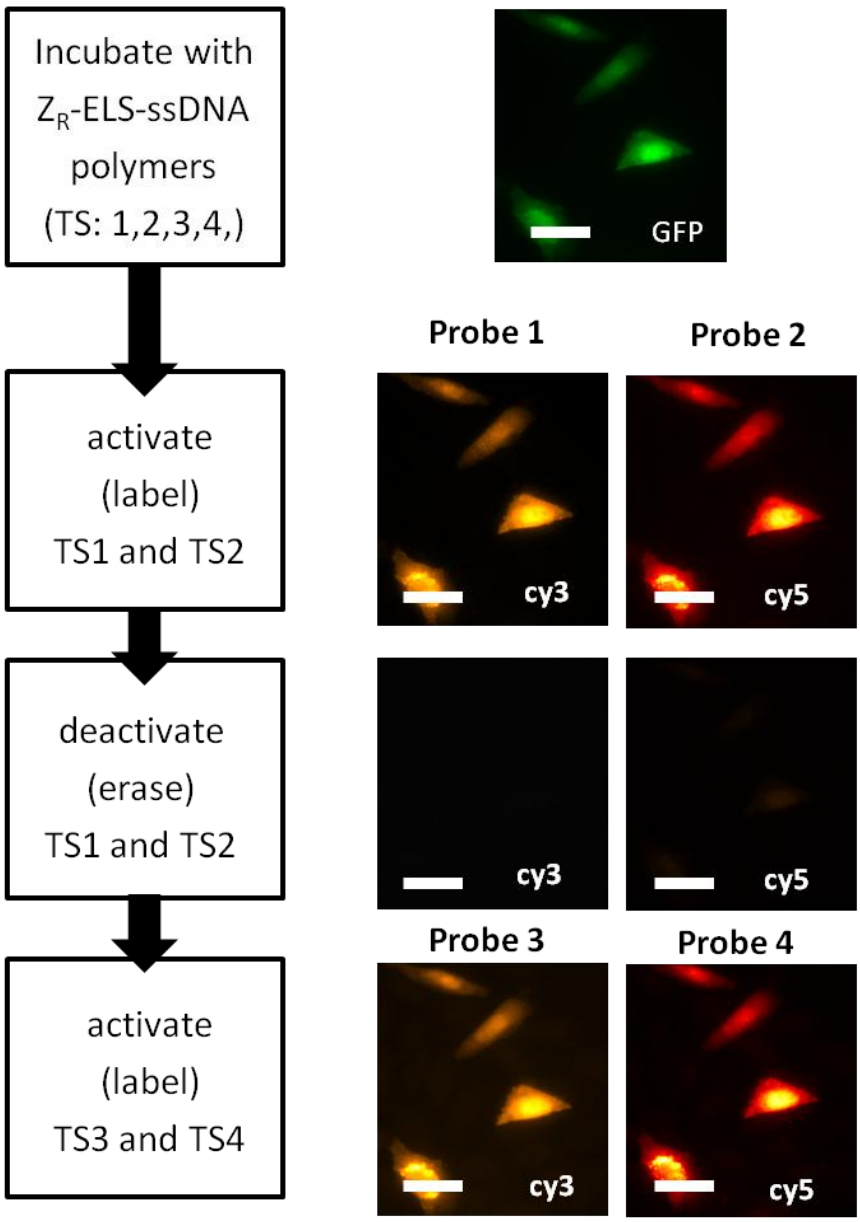

Figure 3.9 Multiplexed (multi-color) and reiterative (sequential) labeling of four different TS strands coupled to the GFP-Z_E targets at equimolar ratios. The scheme at the left depicts the labeling steps for each set of probe images. (Scale bars:50 μm)

3.4. Conclusions

The ability to control the exchange of nucleotides among different dynamic DNA complexes offers opportunities to create new classes of molecular-cell detection and imaging probes that are reconfigurable, adaptive, and that can perform complex logic functions. This detailed characterization of strand displacement reactions within cellular environments allowed the optimization of probe designs for future applications. By comparing the efficiencies and rates that different dynamic DNA -complexes can exchange strands within fixed cell samples as they undergo chemical reactions that either selectively couple (label) or remove (erase) oligonucleotides bearing fluorescent dye molecules to and from DNA-tagged proteins.

This work suggests that while our original 3-strand probe complexes effectively behaved as reiterative labeling complexes, they may not be the optimal design. In addition, half-complexes such as the two-strand complexes are not sufficient when simple three-way strand exchange mechanisms are employed. Therefore, small two-strand complexes which undergo four-way branch migration reactions are optimal for the reiterative labeling of fixed cell samples.

Chapter 4

Multiplexed in situ Immunofluorescence via Dynamic DNA Complexes

Our past efforts effectively characterized and optimized the reactive DNA probe designs for *in situ* molecular analyses. We hypothesized that our probes should, therefore, be capable of being applied directly to immunofluorescence imaging techniques. Using target strand-labeled secondary antibodies we wanted to: 1. Recapitulate images generated using dye-labeled secondary antibodies, 2. Show the ability to reiteratively label and erase cellular markers with a single fluorophore, and 3. Multiplex this capability and demonstrate compatibility with standard immunofluorescence reagents.

To accomplish this, we describe a method to reproducibly conjugate target strands to secondary antibodies, which retain their ability to recognize their antigen. We then stained the microtubules and demonstrated to erase signals to < 10% of their original

intensity, allowing a second marker to be activated using the same fluorescent channel. Furthermore, we multiplex this ability to three marker detection, erase the signals, and detect a second set of markers using fluorophore labeled primary antibodies and phalloidin.

Furthermore, these DNA probes are a platform imaging technology that could be easily incorporated into more complex fluorescence microscopy settings such as super-resolution [102]. Finally, these activation and erasing reactions represent some of the most basic functions that can be performed through strand exchange mechanisms. DNA-based materials have been used to self-assemble complex structures [103] as well as perform logic [60], indicating future potential for strand exchange mechanisms in various immunoassays.

4.1. Introduction

Dynamic DNA complexes are a relatively new class of DNA technologies that can be designed to function as reconfigurable molecular machines [64, 82], detectors [104, 105], programmable logic gates [59, 60, 106], and chemical amplifiers [63, 68, 71]. A unique feature of these devices is that, instead of thermodynamically-driven DNA hybridization reactions, they can harness a process called DNA strand displacement to facilitate the exchange of oligonucleotide strands among different thermodynamically stable DNA complexes to gain kinetic control over the assembly and disassembly of intermediate and product complexes of a chemical reaction [85]. While improved understanding of these processes now provides unique opportunities to program chemical reactions networks for molecular computing applications [65-67], a number of important

biological applications for these devices have also emerged. Dynamic nucleic acid devices have been adapted for *in situ* hybridization analyses of mRNA expression [72], and designed to function 'smart' therapeutic devises that can trigger cell death in the response to the detection of disease-associated transcripts [96]. The use of DNA complexes as molecular delivery vehicles has also been explored [107]. Overall, such advances suggest dynamic oligonucleotide complexes can be engineered to function robustly within complex cellular environments, and can offer abilities to both increase the number and complexity of measurements that can be performed in an assay and introduce new computing functionalities that are not possible using existing probe technologies.

Herein, we demonstrate that dynamic DNA complexes can react selectively with DNA-conjugated antibodies to facilitate multiplexed (*multi-color*) and reiterative (*repetitive*) *in situ* immunofluorescence analyses of proteins within individual cells. For this application, the strand displacement mechanism is used both to create thermodynamically-stable fluorescent reporting complexes that localize to their respective protein target, and to disassemble these complexes so their dye molecules can be removed from cells and used in separate displacement reactions to label additional proteins (Figure 4.1). This capability is important for various molecular-cell studies since it provides a simple, non-destructive, route to increase the number of molecular pathway components that can be visualized within the same cells using fluorescence microscopy. Such analyses are typically limited by the spectral overlap of the fluorophores used for immuno-staining, and generic inability to remove fluorescent antibodies from a sample without employing harsh chemical reagents that can perturb cell morphology and marker antigenicity. Hyperspectral imaging approaches are available that can roughly double the

number of markers that can be imaged simultaneously over conventional methods [108], but further increases have been minimal due to increasing noise and decreasing dynamic range that typically accompanies the integration of additional dye molecules into an immunofluorescence assay [109, 110]. Yet, much more comprehensive proteomic analyses can be performed if cell samples can be labeled, imaged and erased multiple times with or without these advances. Finally, the displacement reactions that are incorporated into the present DNA probe systems constitute elementary components of various programmable chemical networks that have been designed to perform more complex detection functions. Characterizing their ability to interface with immuno-targeting procedures is therefore generically important to the use of such chemical systems for molecular-cell analyses.

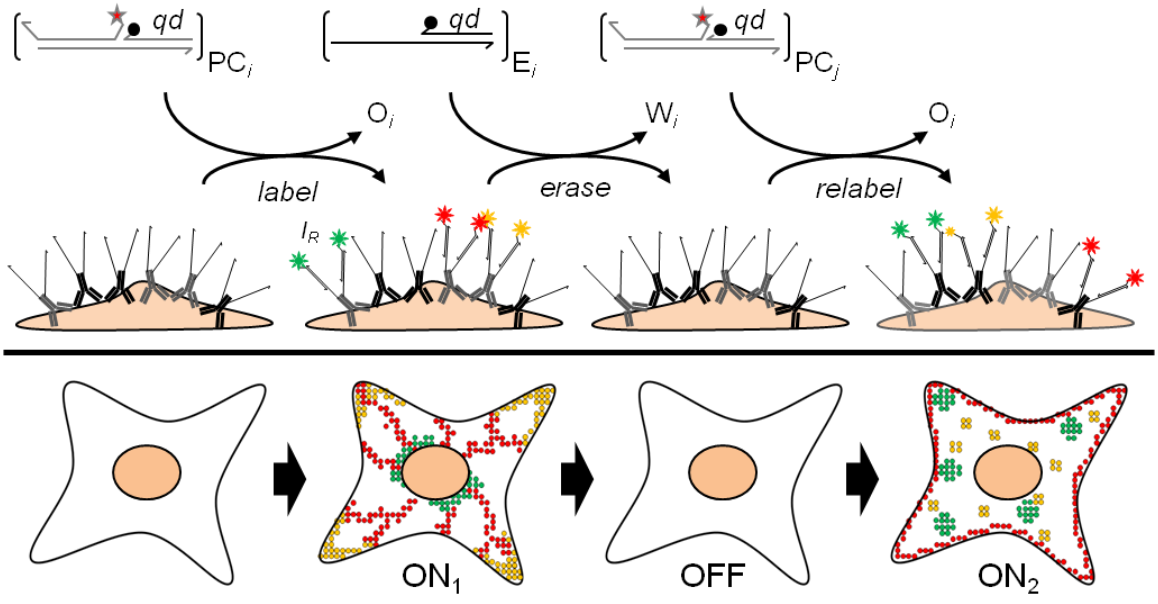

Figure 4.1 Using dynamic DNA complexes, fluorescent markers can be sequentially activated and erased multiple times on a single sample to label multiple sets proteins. Fluorescent signals can be turned ON through the addition of a probe complex (PC) and turned OFF upon incubation with an eraser (E).

4.2. Materials and Methods

4.2.1. General

Goat anti-rabbit (cat# A10533) and anti-rat (cat# A10536) secondary antibodies and phalloidin-Alexa532 were purchased from Invitrogen. Goat anti-chicken secondary antibody, rabbit anti-histone complex H3 (cat# ab1791) and rat anti-α-tubulin (cat# ab6160) were purchased from Abcam (Cambridge, MA). Rabbit anti-KLC4 (cat# HPA030169) and FITC-conjugated mouse anti-Vinculin (cat# F7053) were obtained from Sigma (St. Louis, MO). Alexa647-conjugated mouse anti-WASP (cat# sc-13139 AF647) was purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Oligonucleotides were purchased from IDT, dye and quencher-bearing strands were purchased with HPLC purification. Images were collected with a 14-bit depth CCD camera (Luca, Andor) on a Nikon Ti Eclipse epi-fluorescence microscope with an automated stage and focus, using a 60x oil immersion plan-Apo objective.DNA Circuit Design and Characterization

4.2.2. DNA Probe Design

All sequences were selected using a custom MATLAB script that generates random domains of specified lengths having pre-determined GC% range, while excluding previously generated domains or other prohibitive sequences (*i.e.* G quadruplexes), and avoiding secondary structures (*e.g. hairpins*). The generated domains are ranked according to their normalized two-state hybridization energies with existing probe strands using mFold. The domains are then screened through the BLAST database to minimize

probe sequence homology with the mRNA transcriptome. The final domain sequences are then selected manually from this list and concatenated with other domains to create full oligonucleotide sequences that will be incorporated into a probe complex. Other global criteria such as temperature, strand concentration, and salt concentration are specified prior to domain design. All oligonucleotide sequences can be found in the supplement (Table 4.1).

Strand	**Sequence (5′-3′)**
RQ	GAT GCG AAG TCA GCG TTC/3IAbRQSp/
FQ	CGT AAT AGC GCT AGT CTC/3IABkFQ/
TS1	/5Hexynyl/TTT TTT TTT TGG CCA CCG AGA CAA TAC GCA GGA CCC
TS2	/5Hexynyl/TTT TTT TTT TGT GTA CCG GAA ACA TCG GCG AAT TAG
TS3	/5Hexynyl/TTT TTT TTT TGC CAT CGA CCC GTG CAT TAA GTG TCC
TS4	/5Hexynyl/TTT TTT TTT TAC CTA CGG TCT CCG GAA CTT ACG ATC
D1	/5Cy3/CTT GTC AAT TCG CCG ATG TTT CCG GTA CAC
L1	GTC GGA AAC ATC GGC GAA TTT TTT TTG AGA CTA GCG CTA TTA CG
K1	GTG TAC CGG AAA CAT CGG CGA ATT GAC AAG GAG ACT AGC GCT ATT ACG
D2	/5Cy5/CCT TAA GTC CTG CGT ATT GTC TCG GTG GCC
L2	GTC GAG ACA ATA CGC AGG ACT TTT TTG AAC GCT GAC TTC GCA TC
K2	GGC CAC CGA GAC AAT ACG CAG GAC TTA AGG GAA CGC TGA CTT CGC ATC

Table 4.1 List of oligonucleotide sequences used in design of DNA-circuits. /5Hexynyl/, /5Cy5/, /5Cy3/, /5Alexa488/, /3IAbRQSp/, /3IABkFQ/ indicate a 5′ hexynyl, Cy5, Cy3, Alexa488, 3′ Iowa Black re quencher, or Iowa Black green quencher modification, respectively.

Strand	Sequence (5′-3′)
D3	/5Alex488N/TT TCT CAT AAC ACT TAA TGC ACG GGT CGA TGG C
L3	CTG ACC CGT GCA TTA AGT GTT TTT TTT TTG AGA CTA GCG CTA TTA CG
K3	GCC ATC GAC CCG TGC ATT AAG TGT TAT GAG TTT TTT GAG ACT AGC GCT ATT ACG
D4	/5Cy5/TT TGC AAA GTC GTA AGT TCC GGA GAC CGT AGG T
L4	GTG GTC TCC GGA ACT TAC GAT TTT TTT TTT TTG AAC GCT GAC TTC GCA TC
K4	ACC TAC GGT CTC CGG AAC TTA CGA CTT TGC TTT TTT GAA CGC TGA CTT CGC ATC

Table 4.1 Cont.

4.2.3. Preparation of Antibody-Oligonucleotide Conjugates

Secondary antibodies (20 μg) were incubated with NHS-$(PEO)_4$-Azide (20 μmol in anhydrous DMSO, Thermo Scientific Pittsburgh, PA) in 100 mM $NaHCO_3$ at room temperature for 45 minutes. Following purification via gel filtration using a Tris-buffered Bio-Spin 30 column (Bio-Rad Hercules, CA), hexynyl-modified oligonucleotide were added and a Cu(I) catalyze click reaction was performed using the Click-It Reaction Buffer Kit (Invitrogen) according to the provided protocol. The reaction was incubated at room temperature for two hours then moved to 4º C overnight. The resulting DNA-functionalized antibodies were purified once more using gel filtration. Conjugation was verified with non-reducing SDS-PAGE (Figure 4.2).

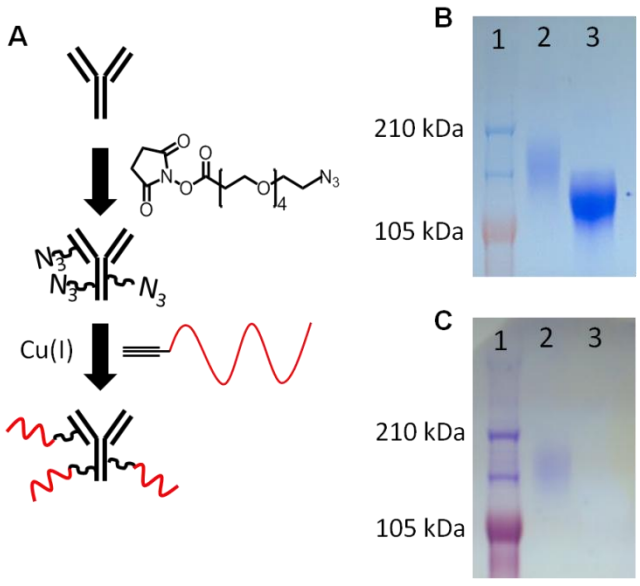

Figure 4.2 Verification of Antibody-DNA conjugation. A. Scheme of labeling reaction using Cu(I) mediated click chemistry. B. Brilliant blue and C. Stains-All stained SDS-PAGE gels of the reaction products: Lane 1: Protein standard, SeeBlue2 (Invitrogen); Lane 2: Reacted DNA-Antibody conjugation product; and Lane 3: Control DNA-Antibody conjugation product.

4.2.4. Cell Culture and Fixation

HeLa cells were cultured in Dulbecco's Modified Eagle Media (DMEM) with GlutaMAX (Invitrogen) supplemented with 10% (v/v) fetal bovine serum (FBS) and 100 U/mL Pen/Strep at 37º C with 5% CO_2. Cells were seeded on cover glass (no. 1.5) and grown to 50-60% confluency prior to fixation. Cells were fixed in freshly prepared 4% (w/v) paraformaldehyde. Samples were then rinsed with PBS and quenched with $NaBH_4$ (1 mg/mL in PBS). The samples were washed with PBS, permeabilized with 0.2% (v/v) TritonX-100, and washed once more. Cover slips were dried under airstream, then affixed to custom-fabricated 10 well chambers using precision cut double sided adhesive and stored in buffer at 4º C for up to one week.

4.2.5. Analysis of Probe Erasing

Samples were blocked 2% (w/v) BSA, 1 mg/mL Herring sperm DNA, and 0.5 μM polyT DNA in PBS, pH 7.4 (buffer BB) for 2 hrs. Rat anti-α-tubulin (1:250) was incubated on samples overnight at 4º C in PBS supplemented with 10% (v/v) buffer BB (PBS-BB). Cells were washed 3 times with PBS, followed by a 2 hr incubation with an oligonucleotide-conjugated (TS1, TS2, or TS3) goat anti-rat secondary antibody at 10 μg/mL in PBS-BB. Cells were then washed 3 times with PBS and probe complex was added to the cells for 2 hours at 100 nM in TAE containing 12.5 mM Mg^{2+} and 10% (v/v) buffer BB (TAE-BB), stained with DAPI (1.3 μM), and washed an additional 3 times with PBS. Erasing complexes were then applied to the cells at 1 μM in TAE-BB, and allowed to react overnight (~12 hrs) at room temperature. Cells were stained again with DAPI then washed 3 times with PBS. Images were collected between reactions steps. After alignment and background subtraction, initial ON images was used to generate a thresholded mask, eliminating pixels with intensity less than 2500. The ON/OFF ratios of these regions (> 400 regions were generated for each image) was then measured directly by comparing pixel intensities of the initial ON values to the OFF images.

4.2.6. Two Marker/One Dye Imaging

Samples were treated as in erasing experiments with minor modifications. Cells were blocked with buffer BB and incubated overnight with antibodies towards α-Tubulin and Histone H3 or α-Tubulin and KLC4 in PBS-BB. Samples were washed with PBS, then TS1 conjugated goat anti-rat and TS4 conjugated goat anti-rabbit were added to the samples at 10 μg/mL in PBS-BB for 2 hours. α-Tubulin was activated with 100 nM

probe complex and erased with 1 μM eraser complex in TAE-BB followed by the activation of the second marker (100 nM probe complex). Images were taken in between reaction steps. Off images were taken after a 10 second bleaching step. All images were aligned as before, false colored, and merged.

4.2.7. Six Marker Imaging

Samples were blocked as before and incubated overnight with rat anti-α-tubulin, rabbit anti-Stathmin 1 (1:200), and chicken anti-Vementin (1:200) at 4º C in PBS-BB. Following washing, cells are incubated with TS1-conjugated goat anti-rat, TS2-conjugated goat anti-chicken, and TS3-conjugated goat anti-rabbit at 10 μg/mL each in PBS-BB. All markers were activated simultaneously by adding 100 nM of each probe complex in TAE-BB. Erasing was performed overnight followed by incubation with 1 μM of each eraser complex in TAE-BB. Alexa647-conjugated mouse anti-WASP (1:500) and FITC-conjugated mouse anti-Vinculin (1:500) were then incubated on the cells for two hours in PBS-BB. Cells were stained with DAPI and Phalloidin-532 then washed. Images were taken between all reaction steps and aligned as before prior to being false colored and merged.

4.3. Results and Discussion

Our antibody labeling reactions exploit ‘toehold domains’ within dynamic DNA probes to initiate strand-displacement reactions between a ssDNA targeting strand (TS) that is conjugated directly to antibodies and a probe complex (PC) that contains a quenched fluorophore (Figure 4.1). These reactions result in the formation of a

fluorescently active reporting complex (I_R) composed of single DNA duplexes. Similarly, a toehold within the reporting complex is used to initiate a second displacement reaction between this complex and an eraser complex (E). In this reaction, the reporting complex is disassembled and its fluorophore bearing strand is rendered inactive by the formation of a waste complex (W) that incorporates a quencher molecule. Consequently, the complete probe labeling and erasing cycle leaves the TS oligonucleotide that is conjugated to the antibody in its original ssDNA state.

Harnessing strand displacement reactions for multiplexed immunofluorescence imaging requires that the DNA-antibody conjugates retain their selectivity for their protein targets and that the coupling and removal of dyes is efficient enough to generate fluorescent images that accurately reflect a protein's intracellular distribution and to ensure these signals do not compromise subsequent imaging analyses of the same cells. These properties were first tested by staining microtubule networks of fixed and permeabilized HeLa cells using primary antibodies raised against α-tubulin, a component of microtubule filaments. Different cell samples were then labeled with one of three unique targeting stands that were conjugated to secondary antibodies and that react selectively with their corresponding PC and E complexes. Stained microtubule filaments are clearly observed in epi-fluorescence images of each DNA-labeled sample (Figure 4.3). The DNA probes also produce microtubule signal to background ratios that were near-identical to those generated by standard dye-conjugated secondary antibodies (*these ratios varied between 10:1 and 20:1 and depended on the region inspected using each method,* Figure 4.4). Similar capabilities were also demonstrated using these probe complexes and an array of primary antibodies that recognize different proteins and

localize to different cellular compartments (Figure 4.5). In each case, the subcellular localization and punctuate staining patterns found using standard immunofluorescence staining procedures are reproduced by the dynamic DNA reactions. Overall, the two methods are primarily distinguished by their ability to resolve fine structures within cell nuclei since non-specifically DNA-complex binding appears to influence the images within these regions of cells, an issue may simply require further optimization of our labeling procedures. These results therefore suggest the DNA-antibody conjugates retain their ability to recognize their protein antigens, and that images capturing the sub-cellular spatial distributions of proteins can be obtained using dynamic DNA complexes to transiently label these targeting agents.

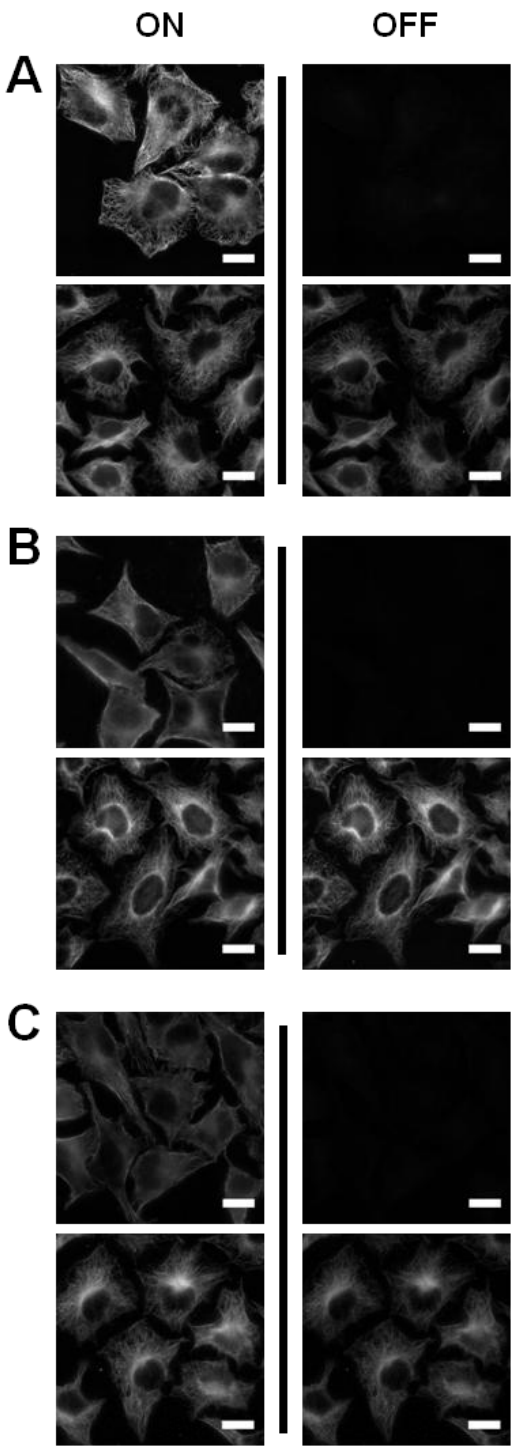

Figure 4.3 Images show the initial "ON" state for A. PS1, B. PS2, and C. PS3 followed by images of cells erased (top) or unerased (bottom). (Scale bars: 20 μm)

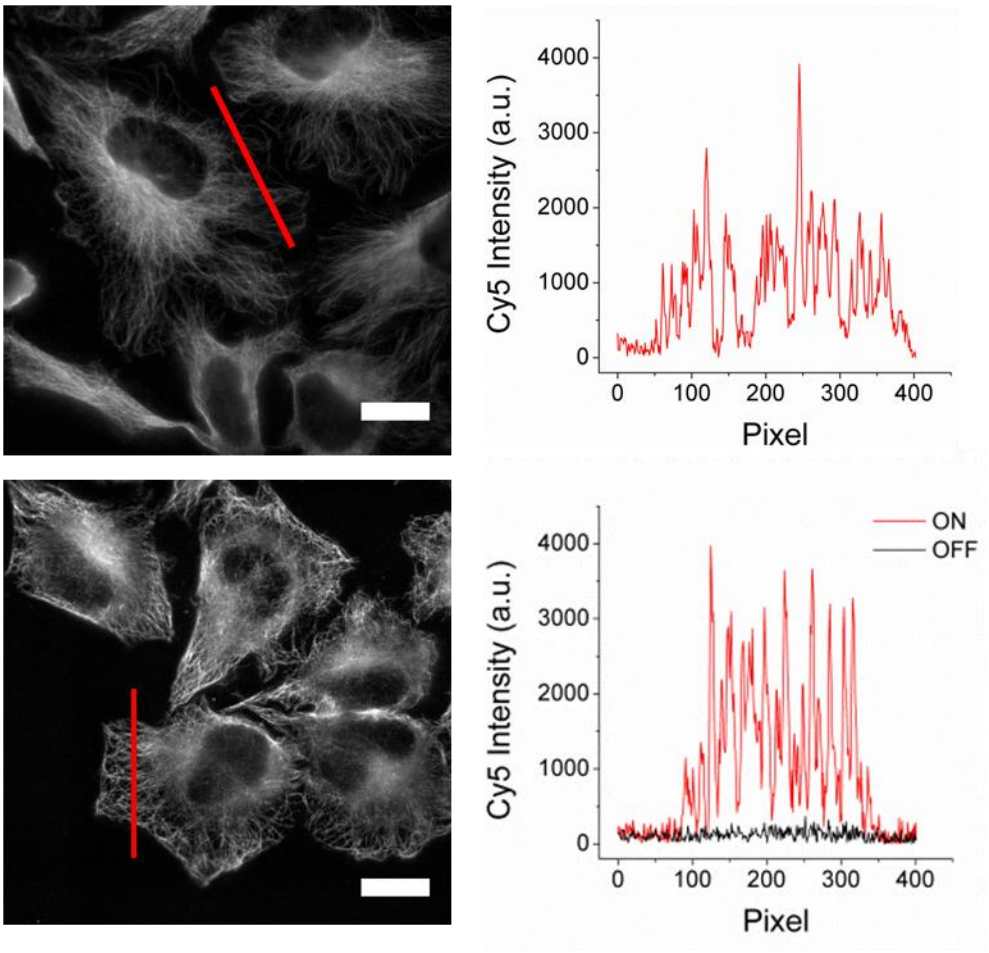

Figure 4.4 Line Profile analysis of α-tubulin detected with a dye-conjugated secondary antibody, signal to noise varies from 1:10 – 1:20 depending upon the location of the filament (top). A similar response in elucidated using DNA-conjugated antibodies detected with DNA probe complexes (bottom). (Scale bars: 20 μm)

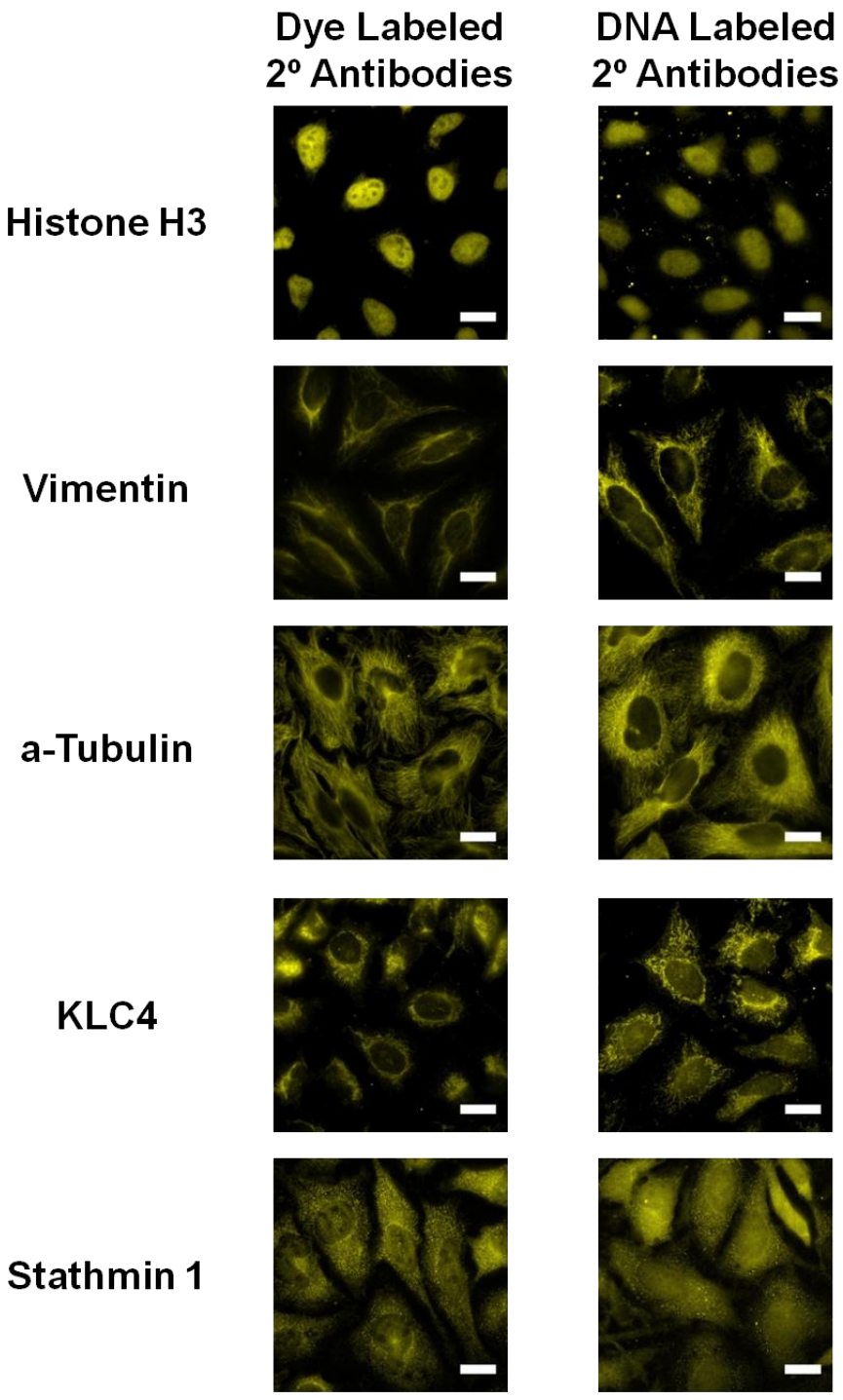

Figure 4.5 (left) Characteristic images of Histone H3, Vimentin, α-Tubulin, KLC4, and Stathmin 1 detected with dye-conjugated secondary antibodies. (right) The same markers were detected using with DNA probes using DNA-conjugated secondary antibodies. (Scale bars: 20 μm)

Strand displacement can also facilitate efficient removal of immunofluorescent signals from previously stained microtubules, yielding ON/OFF ratios between labeled and erased images ranging from 20.0/1 to 24.7/1 (Figure 4.6). At this erasing level, the residual signals are less than or equal to measured RMS fluctuations of background signals within the bare glass portions of the cell culture slides. In contrast, fluorescence intensities are largely unchanged for cells that were labeled using the probe complexes, and then subjected to near-identical incubation and washing procedures used in our erasing reaction, but where the eraser complex is omitted. In this case, signal ratios for sequential images of the same cells varied from 1.36/1 to 1.64/1 (Figure 4.6).

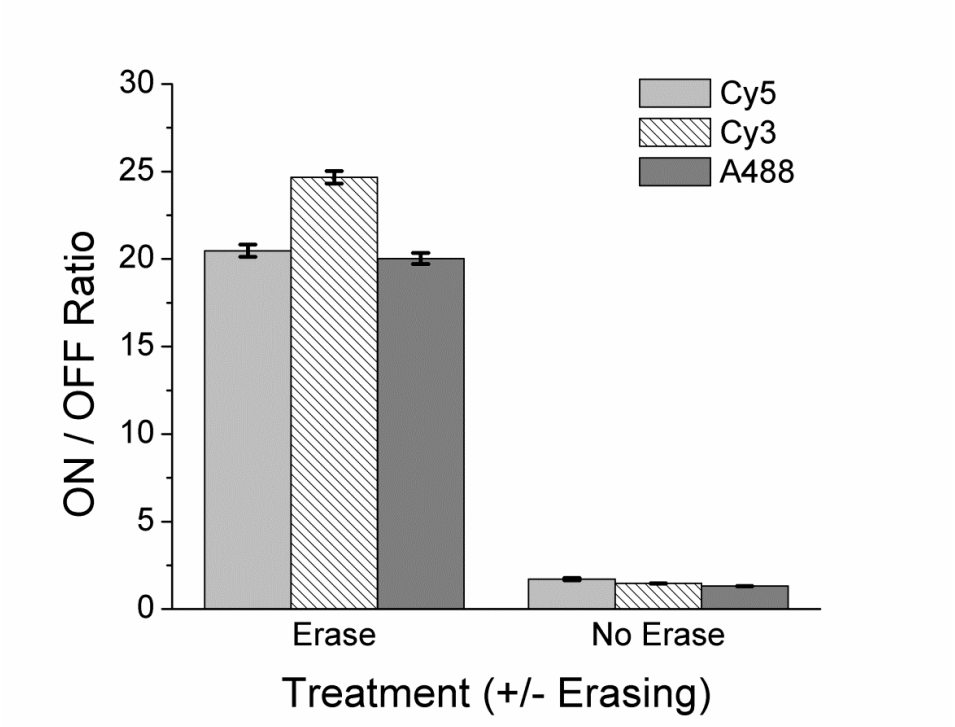

Figure 4.6 Measured ON/OFF ratios of each probe system under all treatment regimens. Erasing alone generated roughly 15:1 ON/OFF ratios for all probe systems (+), but when eraser complexes were omitted, > 70% of the initial signal remained (-).

Prior kinetic studies of analogous *in situ* erasing reactions, where strand displacement was used to remove fluorophores from over-expressed GFP proteins that were tagged with a engineered ssDNA-protein polymer conjugate instead of the current DNA-antibody conjugates, suggest that efficient erasing (*complete and rapid*) requires the use of multi-strand complexes that allow W complexes produced in a displacement reactions to diffuse rapidly beyond a cell's boundary; hence, reducing the probability that reverse / relabeling reactions between W and TS occur. Such control can be gained using three-strand complexes and E strands that liberate two different strands from the I_R, or by employing two-strand eraser complexes that exchange strands with I_R via a four-way branched migration mechanism [111]. However, to address the increasing cost associated with performing multiple immuno-fluorescence measurements, the present DNA complexes were designed to incorporate a quencher domain (*qd*) whose sequence is conserved in each PC and E complex so that the same quenching strand can be employed for all labeling and erasing reactions. While this domain does not participate in the displacement reaction, it precludes the use of entropically-driven, three-strand probe complexes for erasing and is found to reduce the ability to erase probe signals via a four-way branched migration mechanism, dependent upon the incorporated dye molecule (Figure 4.7). This deficiency is potentially due to steric constraints that limit rates the four-way branched migration reactions are initiated since the DNA is directly conjugated to the antibody as opposed to the highly-flexible artificial proteins polymer for our prior analyses, and since eraser complex in this case contains internal toeholds. Nevertheless, this issue was avoided in the above experiments by employing the simple two-strand E complexes depicted in Figure 4.1 that exchange via a three-way branched migration

reaction, and by allowing the displacement reactions with I_R to proceed overnight. Yet, we note that faster erasing could likely be achieving using complexes that avoid the use of the conserved quencher domains.

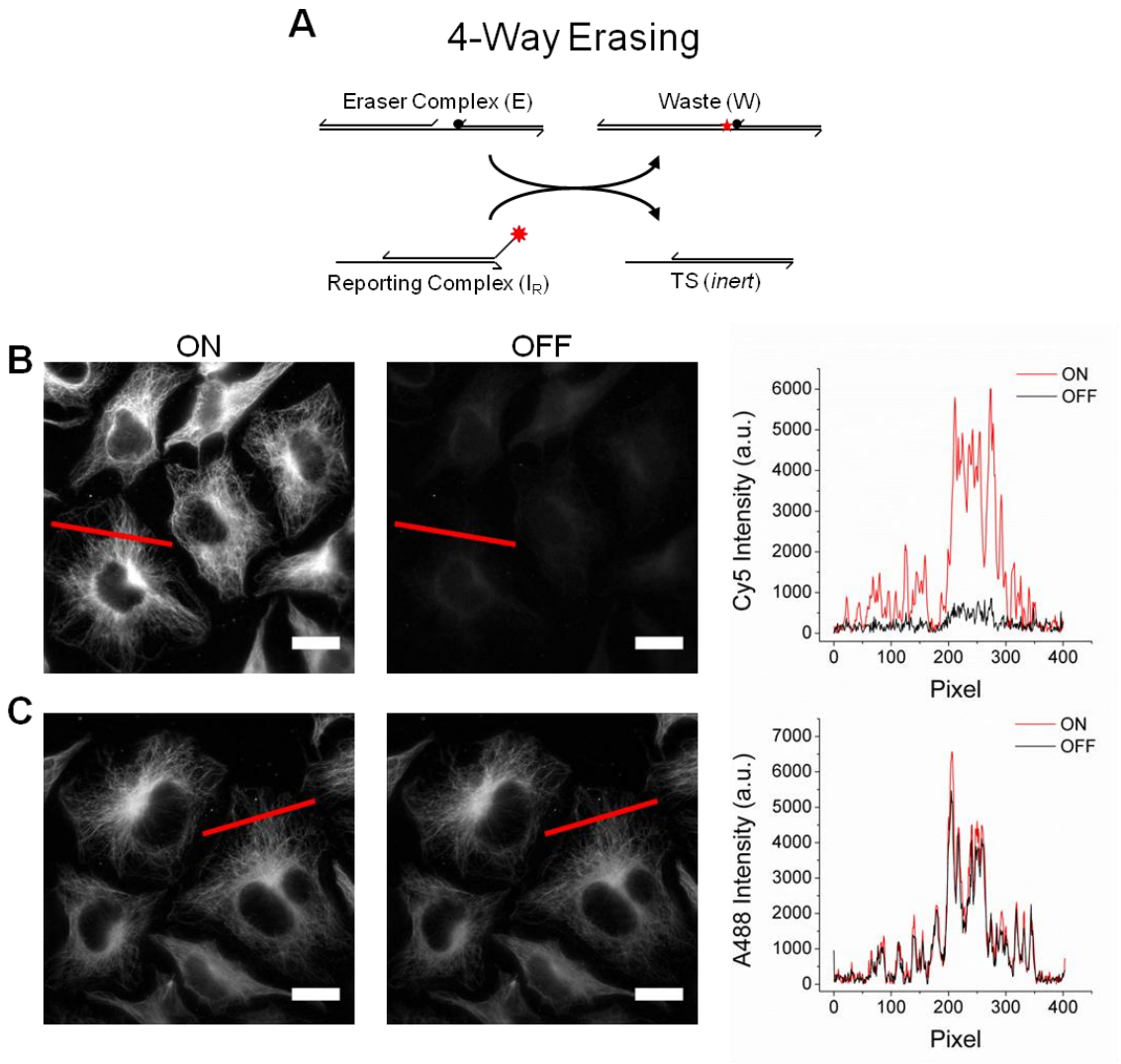

Figure 4.7 A. Scheme for the four-way erasing reaction employing the conserved quencher domain (*qd*). The four-way erasing can reduce signals imparted by Cy5 (B), but is unable to erase Alexa488 dye signals (C). (Scale bars: 20 μm)

The low residual fluorescence signals remaining on a sample after the erasing reactions suggests this procedure can be used to visualize multiple proteins within individual cells using the same color dye molecule. To further assess this capability, additional cell samples were incubated simultaneously with the α-tubulin antibody and a second primary antibody that recognizes either (i) a light chains of kinesin (KLC4) and co-localizes with mitochondria [112]; or, (ii) a Histone H3 complexes that localizes to the cell nucleus (Figure 4.8A). Each of these markers was outfitted with a unique TS strand using DNA-conjugated secondary antibodies (*goat anti-mouse, and anti-rabbit secondary antibodies*). DNA displacement reactions were then used, sequentially, to couple Cy3 dye molecules to the microtubule networks, erase these signals, and then label the second marker. As shown above, microtubule networks are clearly detected in each case, and are erased efficiently upon incubation with E (ON/OFF > 15:1). Moreover, residual signals were sufficiently low to facilitate a second round of immunofluorescence imaging. Here, image intensity profiles are nearly indistinguishable if the final KLC4 and Histone H3 images in these experiments are processed by subtracting their corresponding erased microtubule images or if they are background corrected assuming a constant, spatially-invariant background intensity signal (Figure 4.8B). These results therefore show multiple markers can be inspected using sequential displacement reactions with minimal crosstalk between the signals produce by each reaction.

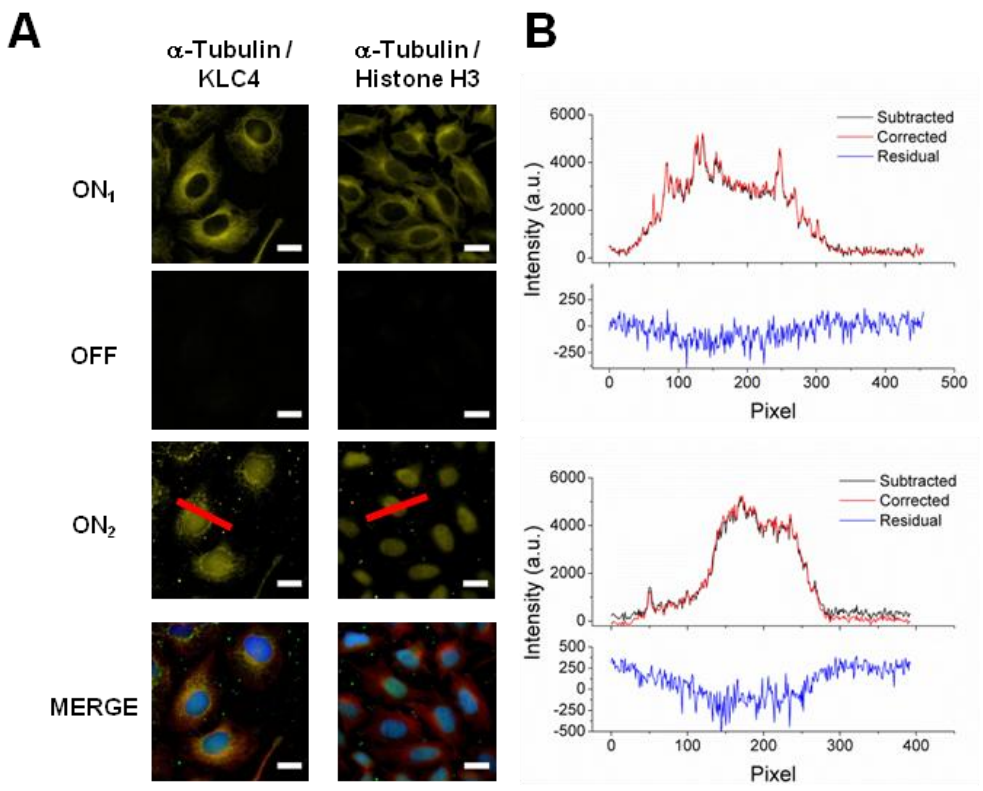

Figure 4.8 (a) Sequential labeling of protein targets using a single fluorescent dye. (b) Line profiles of OFF images subtracted from ON2 plotted against background subtracted ON2 images for cells labeled for α-tubulin and KLC4 (top) and α-tubulin and histone H3 (bottom) with the residual difference between the images shown below (line profile regions are indicated on the ON2 images in (a)).

Finally, we examined the ability to implement our full multiplexed and reiterative imaging procedure to visualize multiple sets of markers with individual cells. For these experiments, six different cytoskeletal-associated proteins **(***stathmin 1, vimentin, α-tubulin, Wiskott-Aldrich syndrome protein or WASP, F-actin, and vinculin*) were imaged, three at a time, using the red, green and blue channels of the fluorescent microscope (Figure 4.9). While the first set of markers were detected and then erased using the three PC or E complexes simultaneously (Figure 4.9; ON_1), the second set of

markers were detected using primary antibodies that were either labeled directly with dye molecules (*Alexa647 conjugated anti-WASP and FITC conjugated anti-vinculin*), or with phalloidin-Alexa532 to stain actin filaments (Figure 4.9, ON_2). Cell nuclei were stained and imaged using DAPI as a forth imaging channel in order to align the images collected in each round of fluorescence microscopy. Again, the signals produce by each DNA probe system reflect the spatial distributions of their protein targets that are obtained using conventional immuno-fluorescence staining methods. Moreover, the ability to stain cells a second time using conventional methods shows that strand displacement can be used to not only double the number of proteins that can be detected within a cell sample, but that antigenicity of proteins targets within cells is also retained throughout this procedure. These results therefore illustrate the flexibility of this approach and suggest the novel detection modalities provided by dynamic DNA complexes can be integrated with various existing immuno-detection technologies in order to custom tailor molecular-cell analyses for specific biological problems.

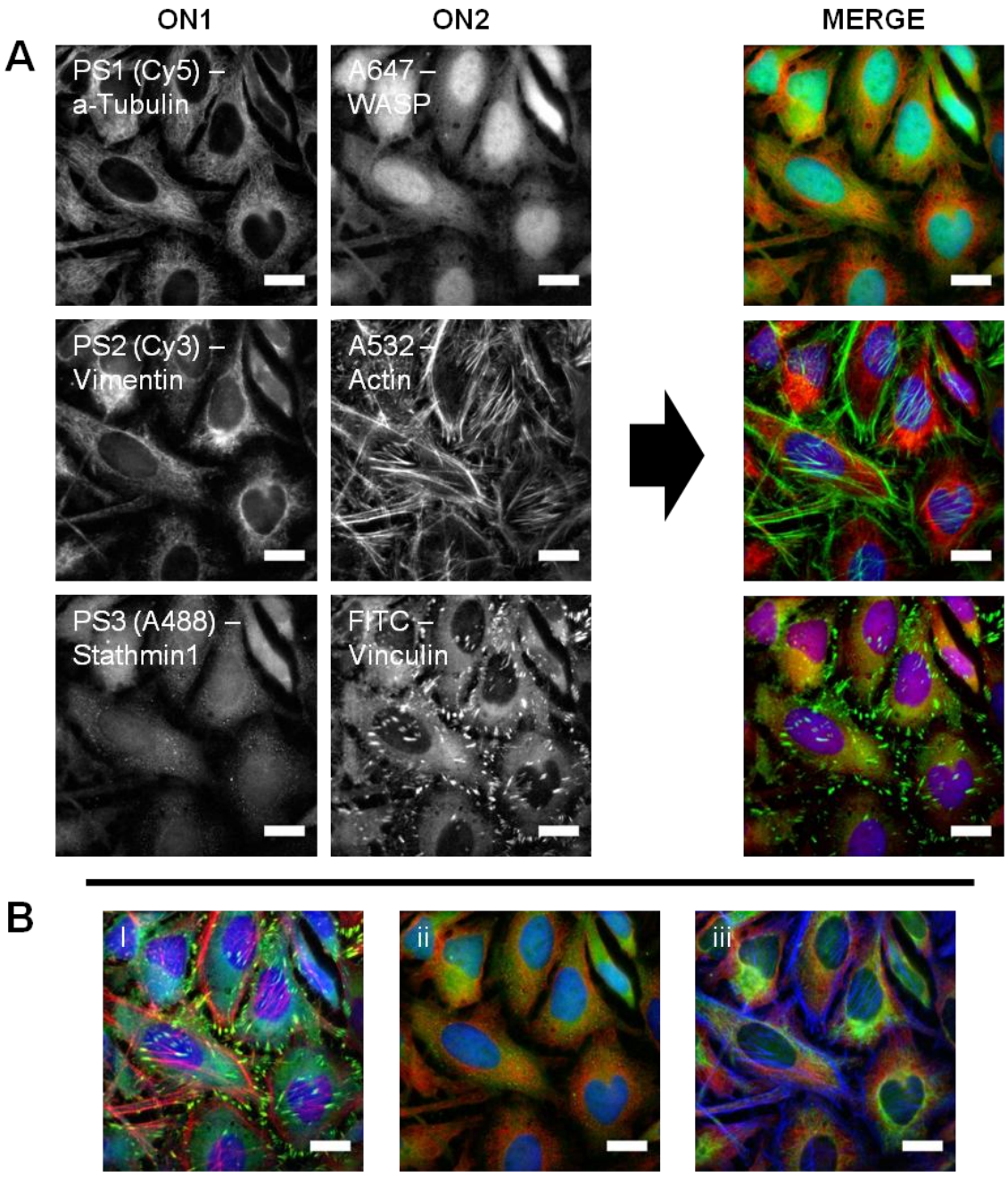

Figure 4.9 Antibodies against α-tubulin, vimentin, and stathmin 1 were detected with DNA-conjugated antibodies labeled by dynamic DNA probes (ON1). The signals were then erased to permit the detection of WASP, vinculin, and f-actin using dye-conjugated primary antibodies and phalloidin (ON2). These markers could then be functionally permuted to examine actin-associated markers (i), microtubule-associated proteins (ii), or cytoskeletal filaments (iii).

4.4. Conclusions

In summary, the present experiments demonstrate the ability to selectively activate and erase fluorescent reporters through the use of strand displacement mechanisms for *in situ* immunofluorescence. Provided steric and kinetic constraints affecting these reactions within cells are addressed, fluorescent dyes can be reused multiple times to at least double the number of detectable markers. Dynamic DNA complexes therefore offer a potential method to label large numbers of protein targets, permitting entire protein pathways to be interrogated on a single sample while retaining any spatially encoded information. Furthermore, as these reactions harness quite basic capabilities of dynamic DNA systems, these complexes may offer new methods to program logic and develop new molecular detection technologies for comprehensive proteomic analyses. This work represents an important first step in achieving these goals.

Chapter 5

A Sequential DNA-Based AND-Gate for Proximity Dependent Fluorescence Detection

Our past work has demonstrated the ability to selectively activate and erase fluorescent signals on fixed cells. These reactions, though, represent basic reaction mechanisms when compared to the dynamic DNA reaction networks described in section 1.5. Therefore, we hypothesized that by modifying our probe complexes to function as AND-gates, proximity dependent protein detection (proximity detection) could be performed.

We altered the probes developed in Chapter 2 to act as sequential AND-gates where two target sequences must be in close proximity and hybridize in a specific order to be detected and produce a fluorescent signal. To test this, we designed a synthetic DNA 'scaffold' that could present two linked targets and analyzed the reactions products

with non-reducing PAGE. A transfectable dimerizing GFP construct was then designed to test the logic gate on cells.

5.1. Introduction

Dynamic DNA-based reactions have been used to create complex logic gates [59, 60] and computational networks [65-67] with specific functions encoded by nucleotide sequence. Importantly, these reactions are highly modular and can be quickly modified and combined to elucidate new functions [63, 68]. In past work, we developed DNA complexes to act as reiterative fluorescence labeling probes to allow elaborate sets of proteins to be visualized within an individual cell or tissue sample [95, 111]. Similarly, the ability of these probes to function as AND-gates could facilitate the detection of co-localized proteins. Such capabilities could be employed to address key limitations of *in situ* molecular imaging for proteomic analyses.

Protein localization and co-localization events can have great influences on protein function as well as cellular phenotypes as a whole [113-115]. The implications of known dimer states have motivated the development of several molecular imaging techniques to visualize these events. While FRET and split-GFP methods have been successfully used to detect protein dimers [116, 117], new direct imaging methods are capable of detecting endogenous co-localization events [55, 56, 79]. Similar to these technologies, our dynamic DNA-based reaction probes should allow the visualization of proximity dependent protein events *in situ*, but the dynamic DNA probes will permit enzyme-free detection.

To test this capability, we have modified our dynamic DNA probe complexes to act as logic elements (AND-gates) for proximity-dependent protein detection. Herein, we describe the design and construction of a sequential AND-gate. This logic gate function was tested using a synthetic DNA scaffold which held two distinct target sequences in close proximity. Lastly, we will demonstrate the utility of these logic gates to detect a dimerizing GFP construct that displays the two unique target strands *in situ*. With this ability, it would be possible to construct a comprehensive library of these logic gate probes to detection several co-localized targets simultaneously and develop new biochemical assays.

5.2. Materials and Methods

5.2.1. General

All chemicals were purchased from Sigma unless otherwise noted. Imaging was performed with a 14-bit depth CCD camera (Luca, Andor) on a Nikon Ti Eclipse epifluorescence microscope with an automated stage and focus, using a 40x air plan-Apo objective. Image processing was performed using ImageJ.

5.2.2. DNA Oligonucleotides

All probe complexes were designed using similar methods to those described in references [58]. Sequences were selected using a custom MATLAB script that generates random domains of specified lengths having pre-determined GC% range, while excluding previously generated domains or other prohibitive sequences (*i.e.* G quadruplexes), and avoiding secondary structures (*e.g. hairpins*). The generated domains are ranked

according to their normalized two-state hybridization energies with existing probe strands using mFold [118], then verified by NUPACK [119, 120].

All DNA oligonucleotides were purchased from Integrated DNA Technologies. Dye or Quencher modified strands were ordered with HPLC purification while unmodified strands were poly-acrylamide gel electrophoresis (PAGE) purified in house prior to use. Complexes were annealed at 20 μM in 10 μL reaction volumes. While dye and quencher complexes were used immediately after annealment, unmodified complexes were PAGE purified, concentrated using ethanol precipitation, and reconstituted in 25 μL of buffer (typically TAE or nuclease-free water). Concentrations were determined from the absorbance measurements at 260 nm. Stocks were prepared at 1 μM concentrations, re-annealed, and stored at -20 C. All oligonucleotide sequences can be found in Table 5.1.

Strand	**Sequence (5′-3′)**
LB	GTA CTG CCT ATA TTC TTC TCA TGA CAA AGT CAA CAT AGT ATA CAG AGG AGT GAT CGA GAT ACT CGA TTA CAA C
OB	GTT GTA ATC GAG TAT CTC CCT CTG TAT ACT ATG TTG AC
SB	TTT GTC ATG AGA AGA ATA TAG G
OutputCC	GTT GTA ATC GAG TAT CTC GAT CAC T
ConsumeCC	GAG ATA CTC GAT TAC AAC
Scaffold1	GCT CCC ACC CGT AGG GAA TGC CCG TTT GCG GTT CCG GGC CGA ACT CG
Scaffold2	CGG GCA TTC CCT ACG GGT GGG AGC TTT CGA GGC GCG TGT CAC GAA AC

Table 5.1 List of oligonucleotide sequences used in for logic gate experiments. /3AmMO/ and /5Cy5/ 3′ amino 5′ Cy5 modification, respectively.

Strand	Sequence (5′-3′)
Scaffold Target1	TCA TGA GAA GAA TAT AGG CAG TAC TTT TTG TTT CGT GAC ACG CGC CTC G
Scaffold Target2	CCT CTG TAT ACT ATG TTG ACT TTG TTT TTC GAG TTC GGC CCG GAA CCG C
Scaffold Target3	ACG ATT CGT TCT AAT AAT AAA GAT TTT TTG TTT CGT GAC ACG CGC CTC G
Scaffold Target4	CGA TTT CTA AGG AAT AAT AAA GAA TTT TTC GAG TTC GGC CCG GAA CCG C
ConsumeCC-3′Q	GAG ATA CTC GAT TAC AAC /3IAbRQSp/
OutputCC-5′Cy5	/5Cy5/GT TGT AAT CGA GTA TCT CGA TCA CT
Amine Target1	TCA TGA GAA GAA TAT AGG CAG TAC TTT TTT TTT T/3AmMO/
Amine Target2	CCT CTG TAT ACT ATG TTG ACT TTG TTT TTT TTT T/3AmMO/

Table 5.1 Cont.

5.2.3. Gel Analysis

For standard logic reactions the Logic Gate, Target Scaffold, and reporter were mixed 1:1:1, 1:1:2, or 1:1:5 (where standard concentrations were 50 nM) in TAE containing 12.5 mM Magnesium. Reactions were performed at room temperature for 3 hours before PAGE analysis. Reaction products were run on 7% (w/v) PAGE gels containing TAE at 150 V for 3.5 hours at room temperature, stained with SYBRGold (Invitrogen), and imaged on an Alpha Innotech (Santa Clara, CA) FluorChem FC2.

5.2.4. Fluorescence Measurements

Fluorescence measurements were performed in Corning (Corning, NY) 96-well plates at 200 μL volumes in TAE with 12.5 mM magnesium on a Tecan (Männedorf, Switzerland) Infinite m1000 plate reader. The target scaffolds and logic gate were

present at 50 nM, while the reporter complex was added at 250 nM concentration. Time zero corresponds to the addition of the logic gate. Measurements were taken every 2.5 minutes for 150 minutes. The excitation and emission were 635 nm with a 5 nm bandwidth and 666 nm with a 10 nm bandwidth, respectively. Readings were taken at 400 Hz flashes with 4 measurements per well. The plate was radially agitated between fluorescence measurements. Measurement settings were optimized to minimize bleaching effects on the fluorescent output.

5.2.5. DNA-Protein Conjugate Synthesis

DNA-Protein conjugates linking the target 1 and target 2 sequences to the $Z_R(ELS)_6$ protein sequence were synthesized as previously described [84]. Briefly, the protein was expressed in BL21 DE3 strain of E. coli and purified. The target sequences were ordered with 3′ amine functionalizations and linked to the protein using a heterobifunctional crosslinker, sulfo-SMCC (Pierce Rockford, IL). The DNA-Protein conjugates were then purified to homogeneity using FPLC. Concentrations were determined using the A260 absorbance of the appended DNA sequence. Conjugates were divided into single-experiment aliquots (generally 10 μL at 10 μM concentration) and stored at -20º C until needed.

5.2.6. Construction of a Dimerizing GFP Protein

The dimerizing GFP protein was constructed using an N-terminal coiled-coil stalk domain of the motor protein Kinesin1 (amino acids 334-401). A GFP and Leucine zipper domain (Z_E) were then fused to the C terminus. The gene was constructed using standard cloning techniques and verified via sequencing.

5.2.7. Cell Culture and Fixation

NIH 3T3 fibroblasts were cultured in Dulbecco's Modified Eagle Media (DMEM) supplemented with 10% (v/v) fetal bovine serum (FBS) with 100 U/mL Pen/Strep. Cells were grown at 37º C with 5% CO_2. Cells were cultured directly on glass coverslips (no. 1.5) to 80% confluence. The cells were transfected with 1.5 μg plasmid DNA/mL media using the FuGene transfection system according to the manufacturer's protocol. After 48 hours, the cells were fixed with freshly prepared 4% (w/v) paraformaldehyde (PFA) in PBS for 30 minutes, washed twice with PBS, then permeabilized with 0.2% (v/v) Triton-X 100 in PBS for 5 minutes. Fixed cover slips were dried under airstream, then affixed to custom-fabricated 10 well chambers using precision cut double sided adhesive. Samples could then be used immediately or stored in buffer at 4º C for up to one week.

5.2.8. Imaging GFP Dimers on Cells

After fixation, cells were blocked for two hours in PBS containing 2% (w/v) BSA, 1 mg/mL denatured Herring sperm DNA, and 0.5 μM polyT DNA. The samples were then washed and incubated with 100 nM DNA-protein conjugates, displaying either Target1 alone, Target2 alone, or both targets in PBS. Cells were washed again and 100 nM logic gate with 500 nM reporter in TAE with 12.5 mM Mg^{2+} was added to each well. Samples were allowed to react for 2 hours, washed, and imaged immediately.

5.3. Results and Discussion

To design the sequential AND logic gate for proximity dependent protein detection, we modified the DNA probe designs from reference [95] (Chapter 2). Using

the ‘catalyst’ strand as the first target and a truncated ‘fuel’ strand (binds at the second toehold, displacing Output2, but not the ‘catalyst’) as the second target, an output should only be generated when both targets are present (Figure 5.1A). The sequential nature of the logic gate comes from the need for target 1 (**T1**) to bind and release the first output strand, which exposes the toehold that allows target 2 (**T2**) to bind. To test the logic gate (**LG**), we created a scaffold which couples the target strands, fixing them at a distance of 24 base pairs (~8.1 nm) (Figure 5.1B). We then incubated target scaffolds, displaying different combinations of correct and incorrect target sequences, with the **LG**. By analyzing the products via PAGE, a full complex is produced when both correct targets are present [Figure 5.1C, **LG** + **(T1 T2)**]. Similarly, when only **T1** is present, the reaction terminates at the creation of an intermediate state [Figure 5.1C, **LG** + **(T1 T4)**]. However, when the **LG** is incubated with only **T2** present, a distinct leakage product is formed (Figure 5.1C, red box). As this indicates a non-sequential binding event, this leakage can lead to false positive results. Therefore, this design was determined to be insufficient for proximity detection.

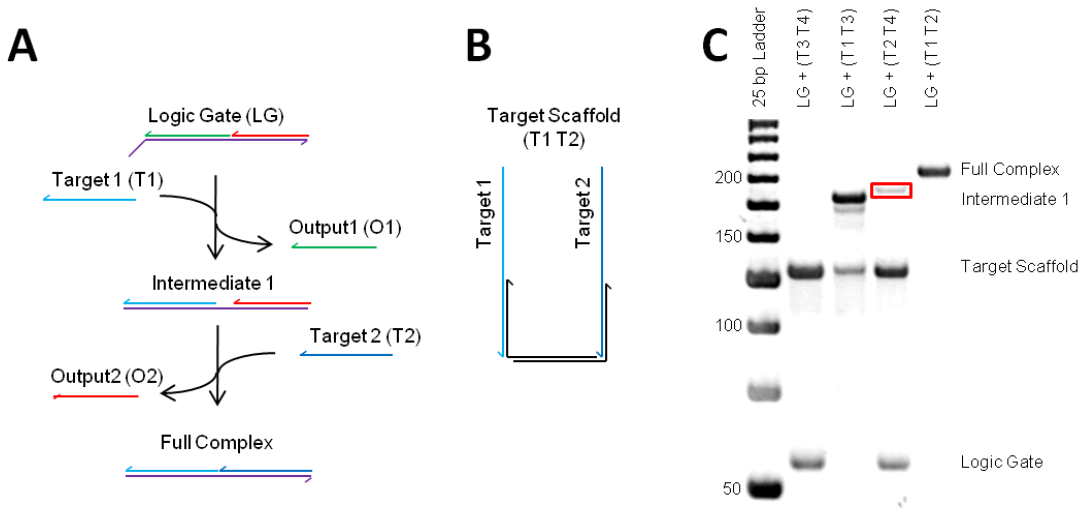

Figure 5.1 Sequential AND-Logic Gate using probe complexes from reference [95]. A. Schematic of the sequential reaction using the entropic circuit design; B. Scematic representation of the target scaffold; and C. PAGE gel depicting the full logic. While a full complex is generated when T1 and T2 are present, a distinct leakage reaction is seen when only T2 is present (red box).

To address the leakage reaction, the second output strand was extended eighteen base pairs to stabilize the complex. This should reduce the ability of **T2** to initiate and hybridize into the **LG** without the presence of **T1**. However, both target strands should be the same length to maintain a modular domain design. Therefore, a seven base pair toehold loop motif was incorporated similarly to reference [90], which is held closed by the second output strand. This design will allow the logic gate to behave sequentially as before, but the full reporting complex can only be formed upon a final reaction with a reporter (**R**) (Figure 5.2A). The new **LG** was incubated with target scaffolds as before, but omitting the reporter complex. Upon PAGE analysis of the products, there is no detectable reaction when incorrect targets sequences are present. When only **T1** is

present, the reaction stops upon the formation of an intermediate state. However, when both **T1** and **T2** are present we see ~70% yield of intermediate 2 (Figure 5.2B). This intermediate 2 state was then incubated with one, two, or five-fold molar excess **R**. Increasing turnover of intermediate 2 to reporter complex is seen as **R** is present in higher excess (Figure 5.2C). Finally, when **LG** is incubated with all versions of the target scaffold and five-fold molar excess **R**, we confirm the sequential AND-logic indicating > 95% yield of reporter complex (Figure 5.2D) by band intensity analysis.

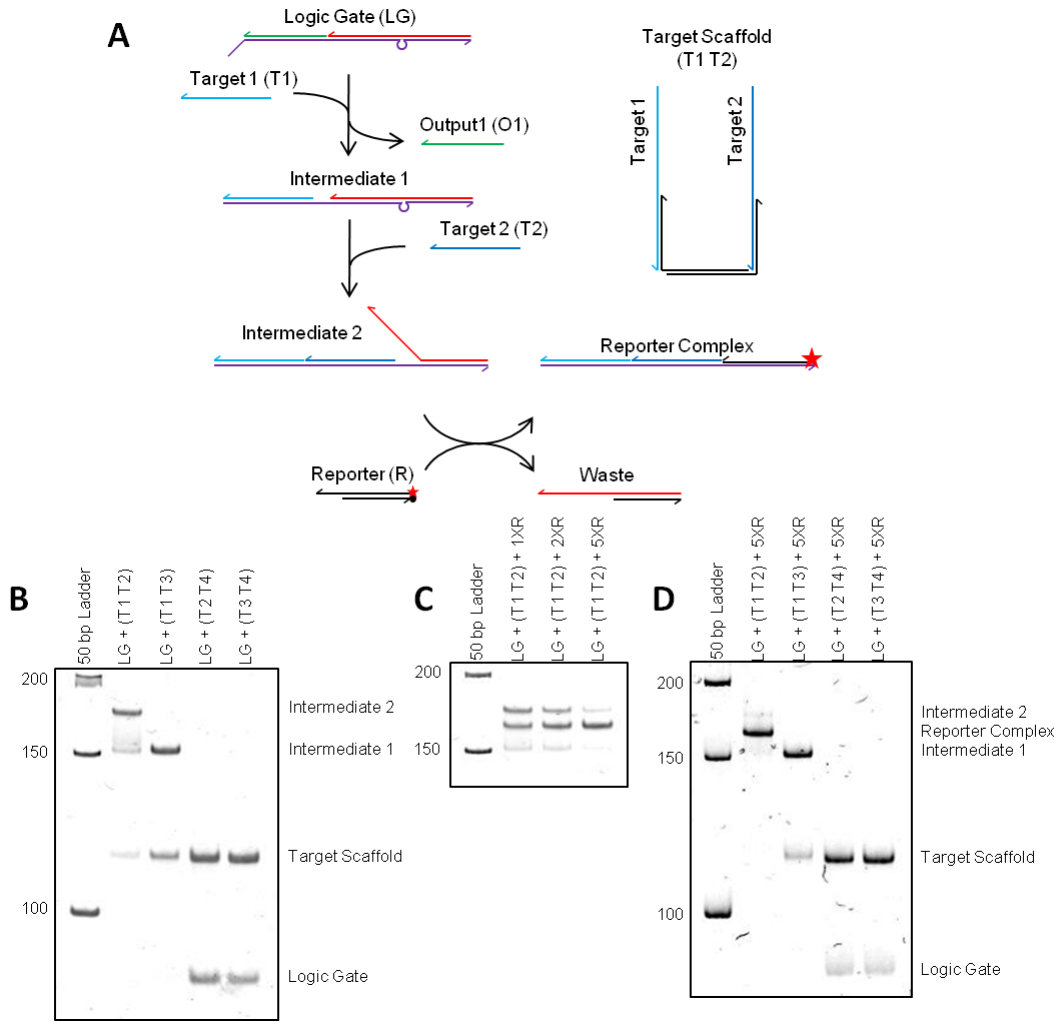

Figure 5.2 Sequential AND-Logic Gate using probe complexes incorporating a new 'loop' domain for proximity detection. A. Schematic of the new logic gate reaction mechanism. B. PAGE gel showing the logic behavior when the reporter is omitted. C. Increasing addition of reporter to the intermediate 2 state drives the reaction to the reporter complex. D. Using 5x excess R, the full logic is driven to completion (reporter complex) only when both T1 and T2 are present.

Next, fluorescent analyses were performed to characterize the solution phase kinetics of the **LG** and precisely quantify background or leakage reactions. By incorporating a dye and quencher into the **R**, the reaction progression can be monitored by measuring the fluorescence intensity. When the reaction was performed using five-

fold excess **R**, equilibrium is reached in approximately 1.25 hours (Figure 5.3). Importantly, a fluorescent output is only detected when the target scaffold containing **T1** and **T2** is added (Figure 5.3, •). Importantly, when **T1** is present alone, the intermediate 1 state is incapable of generating a fluorescent signal (Figure 5.3, •).

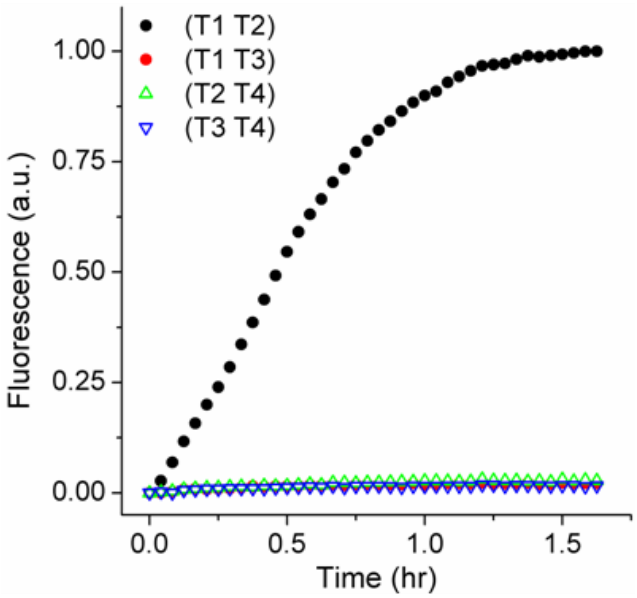

Figure 5.3 Kinetic response a fluorescent reporter (R) when incubated with the logic gate and different target scaffolds. A fluorescent response is seen only when T1 and T2 arepresent (•).

As the sequential AND-gate functionality was verified in solution, we wanted to test the ability of our logic gate to detect protein dimers *in situ*. To do so, we created a homodimerizing GFP fusion protein. By transfecting this construct into cultured cells, we could test our logic gate through the addition of target strand bearing DNA-protein conjugates. GFP-Z_E transfected cells were used as controls as they lacked the dimerizing fusion domain. After adding the DNA-protein conjugates, the cells were washed to remove unbound conjugate and incubated with the **LG** and five-fold excess **R**. Only

when **T1** and **T2** are present do we see a fluorescent response, whereas minimal fluorescence can be detected when one target is omitted (Figure 5.4A). However, a slight fluorescent signal can be discerned from cells transfected with the control GFP-Z_E. We believe this signal is due to the innate weak homodimer formed by GFPs [121].

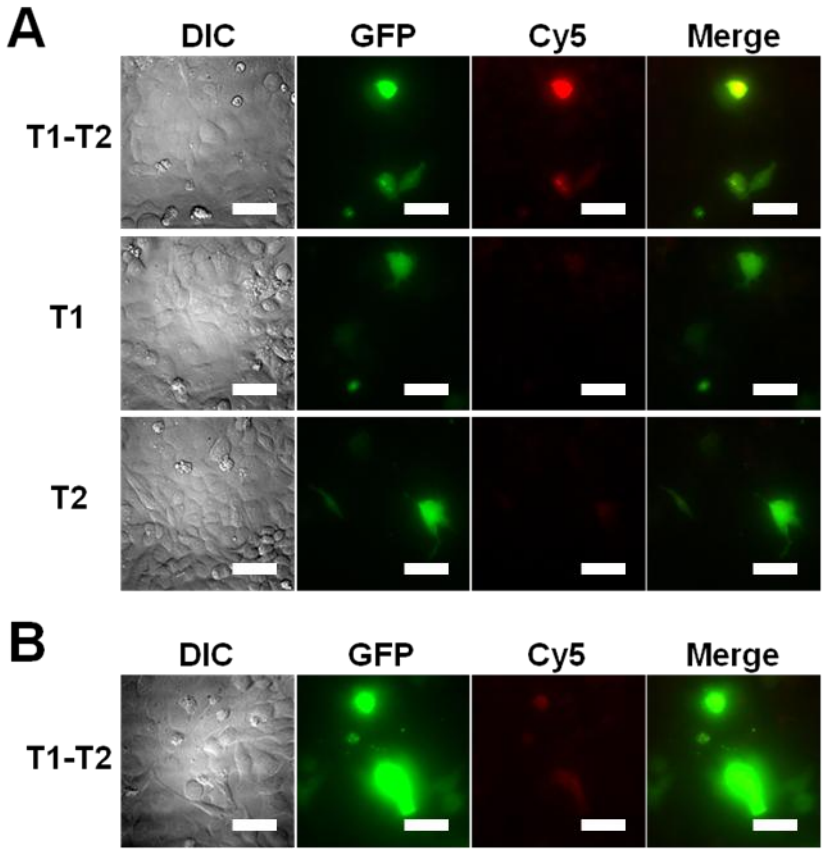

Figure 5.4 Logic gate function on fixed cell samples transfected with a kinesin1 stalk domain followed by a GFP and Z_E. A. The logic gate was added to cells that were incubated with T1 and T2, T1 alone, or T2 alone. Only when both targets are present is a correlation seen between the Cy5 and GFP channels. B. The logic gate was added to cells that were transfected with a GFP-Z_E protein. When both targets are present, a slight response can be seen in the Cy5 channel. We believe this is due to the inherent low-affinity dimerization of GFP proteins. (Scale bars: 50 μm)

5.4. Conclusions

We have shown that the modified dynamic DNA probe complex is capable of function as a sequential AND-gate in solution phase kinetic measurements as well as on cells. The analyses indicate nearly quantitative formation of target complexes through PAGE and show no detectable background or leakage reactions. Similarly, we have demonstrated the ability to directly amplify the detection output using hybridization chain reaction to allow for sensitive detection of low level proximity-dependent events. To fully evaluate the AND-gate probe design, the logic gate should be applied directly to a verified proximity-dependent protein system. Oligonucleotide coupled antibodies or DNA aptamers could then be employed to detect endogenous protein markers and display the appropriate target sequences.

Chapter 6

Conclusions and Future Work

6.1. Summary of Results

In this thesis, we have described the utility of dynamic DNA probes for fluorescence imaging applications. To accomplish this, DNA probe designs were designed and optimized to efficiently and robustly activate and erase fluorescent signals *in situ*. In cellular environments, it was found that molecular crowding can greatly alter dynamic DNA reaction kinetics by enhancing non-toehold mediated strand displacement reactions. Therefore, it was necessary to redesign circuit complexes to either generate inert products or require the reverse reaction to have a cooperative element. Antibodies were then labeled with DNA to employ these probes for immunofluorescence imaging. Here, all probe designs were capable of producing antibody-based signals on fixed cell samples. The dynamic DNA probes were then capable of efficient erasing to allow at least one additional round of labeling and detection. Furthermore, we have utilized the

reconfigurable and highly predictable nature of DNA hybridization reactions, to perform simple AND-logic for proximity dependent protein detection. The modular domain design of the dynamic DNA probes allowed the construction of a sequential AND-gate, which only produced a fluorescent reporter upon the detection of two distinct co-localize target sequences.

We have, thus, established a new class of reconfigurable fluorescent probes that harness strand exchange mechanisms to activate and erase fluorescent markers and detect co-localized proteins *in situ*. The ability to incorporate complex DNA-based computations and logic offers new avenues for molecular imaging. These capabilities may permit more comprehensive proteomic molecular pathway analyses, characterizing the network-level properties of cells. Such advancements have important implications in disease and cancer diagnostics given the ability to specifically characterize both sample heterogeneity as well as rare cell types. Through the creation of stronger functional networks among known and unknown cancer biomarkers new and effective diagnostic assays and treatments may be devised.

6.2. Future Work

6.2.1. Improving Sensitivity and Detection

While Chapter 4 indicated the power of the reiterative technology using cytoskeletal markers, it may be of interest to improve our detection capabilities for low-level expression markers. To improve upon our detection sensitivity, we should optimize our conjugation methods and efficiency and develop a DNA-based method to amplify fluorescent outputs.

Control over the number of covalently attached target strands may offer a way to incorporate a larger number of dyes per bound antibody. The current 'click'-based conjugation relies on a sufficient number of NHS-azide groups to be coupled to the antibody for labeling. While Guo *et al.* suggested that the protocols used should facilitate two to three strands per antibody [58], we have yet to verify these values. Even if we assume these values to hold, we do not know if all conjugated target strands are capable of reacting with the probe complexes due to sterics. Similarly, their oligonucleotides were composed purely of synthetic nucleic acid structures and were much shorter than our target strands, suggesting that the reported conjugation efficiencies could vary greatly when compared to our studies. To address this, a 5′ hexynyl-modified target strand could be ordered with a 3′ dye modification. Using our conjugation methods, the degree of labeling could be easily calculated using simple absorbance measurements. Furthermore, by hybridizing this target strand with an unlabeled oligonucleotide prior to conjugation, the reaction conditions could be optimized to minimize steric interference among target strands.

Alternatively, new conjugation methods could be explored. Multiple groups have reported the covalent attachment of dyes and other molecules to the carbohydrate chains located on the conserved region, Fc, of antibodies using 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) coupling. While optimization methods could be performed identically to the 'click' methods, EDC coupling would also prevent antibody loss due to conjugation at the variable, antigen-recognizing region, Fv.

To amplify signals imparted by the DNA complexes, we are currently proposing a method to create multivalent structures which are initiated by a single target strand. By

constructing "trees" from a single target strand, each "branch" can be recognized individually by a probe complex. Importantly, these trees also can be multiplexed to enhance the degree of amplification (i.e. a single trees can display three additional trees). Simultaneously probe complexes can be designed to incorporate additional dye molecules for additional amplification. Since these trees are finite size structures, the level of amplification can be programmed to specifically modulate a signal for a given application. Currently, these trees have shown the ability to amplify our GFP model system by factors of 3 or 9 (Figure 6.1) via the multiplexing of tree structures or the incorporation of multiple dyes in a single probe complex.

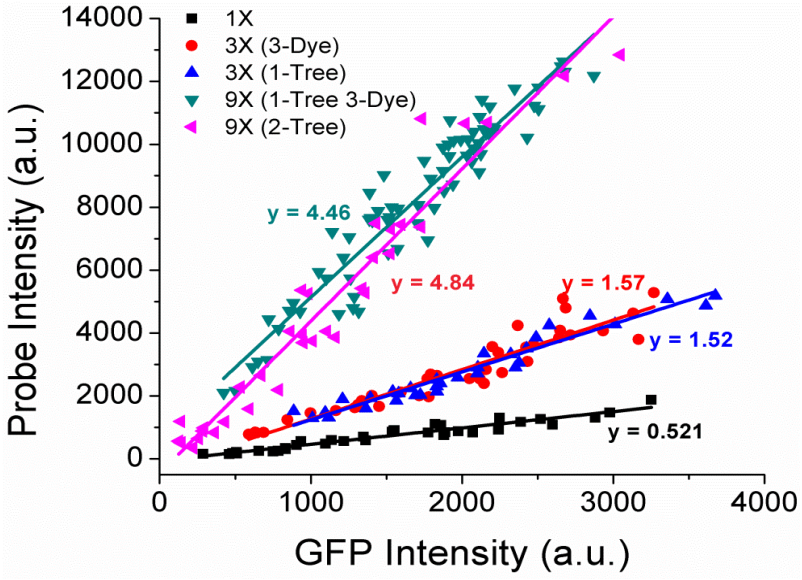

Figure 6.1 Graph of the quantile modulation of dye intensities through "tree"-based amplification. The fitted slopes are indicated on the graph, showing 3x enhancement by incorporating three dyes (•) into the probe complex or using a single tree (▲). Similarly, 9x amplification can be achieved through two levels of trees (◀) or a single tree with probe complexes containing three dyes (▼). Figure courtesy of Jan Zimak.

6.2.2. Application of Logic Gate

In Chapter 5, we demonstrated the ability of our modified DNA-based logic gate to detect co-localized GFP molecules. Having used a synthetic, transfected system, it would be prudent to demonstrate functionality by detecting endogenous protein co-localization events. Given the recent establishment of a commercially available proximity ligation kit (Olink Biosciences), we should be able to validate the logic gate using protein pairs previously verified by PLA. Using these well characterized antibody pairs, we can conjugate secondary antibodies with our target strands and evaluate the ability of the logic gate to detect proximity-dependent co-localization events.

Chapter 7

Design of DNA-Conjugated Polypeptide-based Capture Probes for the Anchoring of Proteins to DNA Matrices

A critical aspect of fluorophore-based imaging is targeting specific cellular markers. Most often, antibodies, small molecules, and biological analogues are employed to facilitate specific detection. As a model system, we believed that transfectable targets would be ideal markers to develop the reactive DNA probe complexes. While several semi-synthetic DNA-protein conjugates have been described in the literature, we wanted to create a monovalent conjugate capable of coupling a single strand of DNA to a single protein.

For this application, we chose heterodimerizing leucine zippers (Z_E and Z_R) due to their high affinity ($K_D \sim 10^{-15}$) and monovalent behavior. The Z_R zipper was linked to a long flexible elastin-like linker terminated with a cysteine residue for attachment

chemistries. Importantly, this protein exhibits no major secondary structure, permitting denaturing expression and purification methods while able to withstand multiple solvent conditions for efficient chemical conjugation. To demonstrate the capabilities of these conjugates, they were used to attach fusion proteins to DNA-addressed locations on microarrays and arrange into finite size structures in near unity efficiencies, all while maintaining protein/enzyme activity.

While the experiments in this chapter detail the use of linking fusion proteins to DNA-based scaffolds, this semi-synthetic conjugate would allow the direct attachment of target DNA strands to transfected fusion proteins similar to those described in these experiments.

7.1. Introduction

Protein coupling and immobilization strategies remain central to the investigation of protein interactions and constitute an important design element for a wide variety of protein detection technologies [122-124]. To meet these needs, an array of polypeptide-based affinity tags has been implemented that can mediate the surface anchoring of proteins [125, 126]. In addition to being expressed as small protein fusions, these

functionalities offer significant advantages over physical adsorption as well as most direct covalent attachment schemes since they provide control over the orientations of proteins on their supports and, in many circumstances, help preserve the native folded state of proteins [127].

Alternatively, DNA-protein conjugate molecules have been developed to facilitate protein immobilization onto DNA-coated matrices such as DNA microarrays [128-130] and organized molecular scaffolds formed from DNA [131]. DNA-directed immobilization of proteins allows oligonucleotide sequences to encode the spatial positioning and composition of multiple proteins on solid supports via the controlled deposition of DNA [128]. This app-roach can simplify the preparation of solid supports for protein immobilization and provide sensitive control over protein deposition since DNA-surface chemistries are more established and robust than most protein immobilization techniques. DNA-directed protein deposition can also provide routes to tune the quantitative activities of enzymes in ways that can be difficult to achieve using protein affinity tags alone and offer new opportunities to build artificial systems of interacting proteins with complex kinetic properties. Such control has been demonstrated by recent work to ‘program’ the stoichiometric ratios of multiple enzymes immobilized on the surfaces of DNA-coated 96-well plates [132]. Similarly, multi-enzyme complexes have been created using DNA as a molecular scaffold and used to explore how proximity effects influence their coupled-enzyme reaction dynamics [132, 133].

Despite the advantages of using DNA hybridization to directed protein deposition, various technical obstacles still limit the utility of this approach for a variety of applications. Particularly, the production of large numbers of DNA-protein conjugates

remains challenging. Although commercially available crosslinking agents can be used to couple amine or thiol-functionalized DNA strands to surface cysteine and lysine residues of proteins [134, 135], the efficiency of these reactions tends to be low and post-labeling purification is generally required to isolate the conjugates. Furthermore, site-specific DNA conjugation through this route requires that the majority of surface cysteines and lysines are mutated to non-reacting residues. Even if these, or alternative modifications are possible, the direct labeling of proteins with DNA may interfere with protein function [135]. These issues can be avoided by coupling DNA to proteins indirectly through protein affinity tags. Protocols that employ intein tags along with expressed protein ligation [136, 137], DNA-streptavidin conjugates and biotin labeling [128], reactions of chemically modified DNA molecules with SNAP-tag fusion [138], and protein farnesyltransferase labeling coupled with click chemistry [139] have all been adopted for this purpose. However, most of these methods still require purification steps to isolate the DNA-protein conjugates. Consequently, a host of technological applications stand to benefit from the development of new synthetic routes to prepare DNA-protein conjugates that emulate the simple and robust processing afforded by commonly used protein affinity tags alone.

7.2. Materials and Methods

7.2.1. General

All reagents were purchased from Sigma-Aldrich unless otherwise noted. All oligonucleotides were purchased from Integrated DNA Technologies. The Alexa647-anti-green fluorescent protein (GFP) was purchased from Molecular Probes (Carlsbad,

CA). Anti-GST antibodies were purchased from GE Healthcare (Piscataway, NJ) and labeled in house using a Cy3 mono-reactive dye pack (GE Healthcare) and the resulting dye to antibody ratio (6 to 1) was determined according to the manufacturer's protocols.

All cloning procedures were carried out in the *E. coli* strain XL1-Blue (Stratagene Wilmington, DE). Expression was performed using BL21 (DE3) strains (Novagen Wilmington, DE). The sequences of the plasmids produced from each cloning step were verified using restriction mapping and DNA sequence analysis.

Primers sequences used to construct synthetic genes were designed using a web-based software tool (DNAWorks) [140]. DNA sequences used for conjugation, surface-functionalization, and to form molecular scaffolds based on single DNA duplicies were designed using a program called SEQUIN [141]. This software was used to minimize sequence repeats in the different strands used for conjugation and spotting, as well as those incorporated into the DNA scaffolds. All oligonucleotides were individually purified using denaturing PAGE protocols and can be found in Table 7.1.

Sequence Name	Nucleic Acid Sequence
Conjugate1	/5AmMC6/CCAATGCGGTCTATCCAGCC
Conjugate2	/5AmMC6/CTACGGCAACTGTGGTCATC
scConjugate1	/5AmMC6/TTTTTTTTTTTTTTTTTTTTGGCTGGATAGACCGCATTCG
scConjugate2	/5AmMC6/TTTTTTTTTTTTTTTTTTTTGATGACCACAGTTGCCGTAG
sControl	/5AmMC6/TTTTTTTTTTTTTTTTTTTTCCGAGCAGGCGTGAGTCGTC
kcConjugate1	/5Biosg/TTTTTTTTTTTTTTTTTTTTGGCTGGATAGACCGCATTCG
kControl	/5Biosg/TTTTTTTTTTTTTTTTTTTTGATGACCACAGTTGCCGTAG
Scaffold Overhang1	CGTAGCAGGCACATCGTTGGCTGGATAGACCGCATTCG
Scaffold Overhang2	CGTAGCAGGCACATCGTTGATGACCACAGTTGCCGTAG
Scaffold1	GTCACGGACTGAGCGT
Scaffold2	CGATGTGCCT
Scaffold3	GCTACGACGCTC
Scaffold4	AGTCCGTGACTTGGCTGGATAGACCGCATTCG

Table 7.1 List of nucleic acid sequences used for conjugation, microarray experiments, and multiprotein scaffold experiments. /5AmMC/ indicate a 5′ amine modification and /5Biosg/ indicate a 5′ biotin modification.

7.2.2. Plasmid Construction

Genes encoding for each domain of the artificial proteins including the basic portion of the zipper complex (Z_R) and the elastic 'mid-block domain' or ELS fragment (VPGVG VPG**S**G VPGVG VPG**S**G VPGVG) were constructed from synthetic oligonucleotides through PCR. During this process a C-terminal cysteine and a stop codon was introduced into the gene. The resulting fragments were directionally ligated into a pQE60 vector (Qiagen Valencia, CA). This gene was then polymerized stepwise

into larger Z_R-$(ELS)_2$-C, Z_R-$(ELS)_4$-C and Z_R-$(ELS)_6$-C constructs using the recursive directional ligation (RDL) procedure described in reference [142] with slight modifications. At each intermediate cloning step, two gene fragments were isolated from each vector by digestion with *Nco*I and *Bgl*I or *PflM*I and *Hind*III restriction endonucleases. These two fragments were then simultaneously inserted directionally into the *Nco*I and *Hind*III sites of a new pQE60 vector. The amino acid sequences for the Z_R-$(ELS)_6$-C protein can be found in Table 7.2.

The recombinant target proteins containing Z_E fusions GST-Z_E, GFP-Z_E, calmodulin-Z_E (CAM-Z_E), and PSE-4-Z_E were all prepared using standard cloning procedures. Each of these genes contained a C-terminal Z_E and 6xHistidine (6xHis) for purification.

Sequence Name	**Amino Acid Sequence**
Z_E	LEIEAAALEQENTALETEVAELEQEVQRLENIVSQYRTRYGPL
Z_R	LEIRAAALRRRNTALRTRVAELRQRVQRLRNEVSQYETRYGPL
ELS	VPGVGVPGSGVPGVGVPGSGVPGVGV
Z_R-$(ELS)_6$-C	MVGS-Z_R-GSNHGVG-$(ELS)_6$-VPGWLRSC

Table 7.2 Amino acid sequences of the artificial polypeptide components in the Z_R-$(ELS)_6$-C protein.

7.2.3. Protein Expression and Purification

To avoid proteolytic degradation of the Z_R fragment of the artificial proteins, and to facilitate Ni^{2+}-NTA purification, Z_R-$(ELS)_6$-C was co-expressed with a Z_E-6xHis gene that was created in pQE60 and then inserted into the *Nhe*I sites of a pREP4 vector

(Qiagen). Both vectors contained the inducible, bacteriophage T5 promoter of the pQE plasmid. For expression, cells were grown in terrific broth (TB; 24 g of casein hydrolysate, 12 g of yeast extract, 4 mL glycerol, 17 mM KH_2PO_4, 72 mM K_2HPO_4), and antibiotics (50 mg/L ampicillin and 35 mg/L kanamycin). The cells were cultured at 37 C until an OD_{600} of 1 was reached. Afterwards, expression was induced for 5 hours by adding 1 mM IPTG. Cells were then pelleted (5,000 RPM and 4 C) an resuspended in lysis buffer (8 M urea, 10 mM Tris-HCl, 100 mM NaH_2PO_4, pH 8.0) and allowed to lyse overnight at 4º C with vigorous stirring. The cell lysate was clarified by centrifugation and the supernatant was applied to a column containing Ni^{2+}-NTA resin. After washing with lysis buffer, the artificial protein was eluted using a buffer containing 6 M guanidine-HCl, 100 mM NaH_2PO_4, 10 mM Tris-HCl, pH 8.0. This buffer disassociates the Z_E/Z_R complex, allowing the Z_R-$(ELS)_6$-C to be eluted while the majority of the Z_E-6xHis proteins remain on the column. To remove excess Z_E-6xHis, the eluent was run through a second Ni^{2+}-NTA column. The flow through of this column was dialyzed with frequent water exchanges for 3 days and lyophilized. The protein purity was verified using SDS-PAGE and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectroscopy (Figure 7.1). The MALDI-TOF of the polymers gave peaks at 18652 m/z and 9323 m/z corresponding to the M^+ and M^{+2} polymer species, respectively. The Z_E-6xHis polypeptide was not present at detectable levels using either SDS-PAGE or MALDI-TOF analyses.

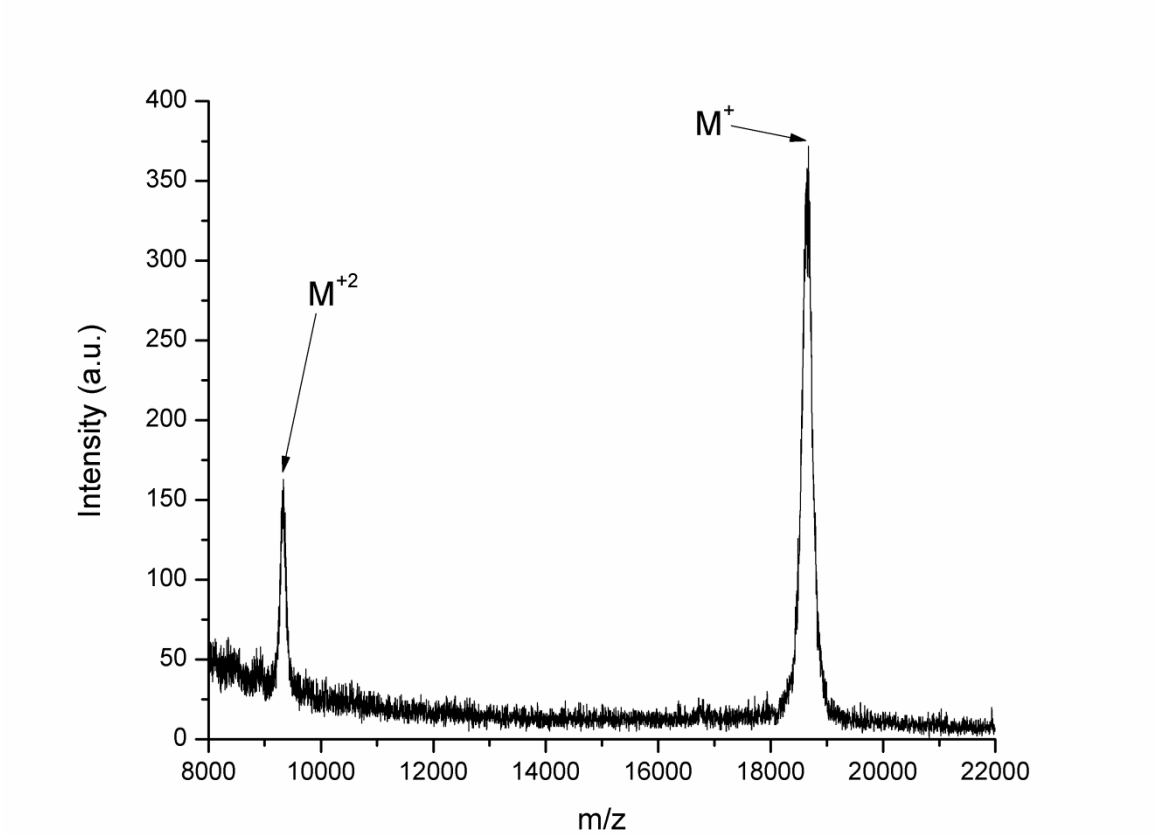

Figure 7.1 MALDI-TOF spectrum for the Z_R-$(ELS)_6$-C protein. The spectrum shows peaks at 9323 m/z and 18652 m/z, corresponding to the M+2 and M+ peaks, respectively. The calculated theoretical mass of the full-length polymers is 18793.5 Da, and 18662.3 Da if the N-terminal methionine is cleaved.

Cells expressing the target proteins were cultured similarly to those used for artificial protein production. However, expression was induced overnight at 23 C. Purification of the target proteins were carried out using standard native lysis and Ni^{2+}-NTA purification protocols with the exception of the PSE-4-Z_E. Cells containing these proteins lysed using an osmotic shock procedure that allows the outer membrane of cells to be broken using an osmotic shock buffer [30% (w/v) sucrose, 1 mM EDTA, 30 mM Tris, pH 8.0]. Protein expression and purification was verified using SDS-PAGE

analysis while protein concentrations were determined using the Bradford method. For some experiments, clarified lysates were saved, aliquoted, and stored at -80º C for experiments with unpurified proteins.

7.2.4. Artificial Protein-DNA Conjugation

Artificial proteins were labeled with DNA as described in reference [134] with slight modifications. Briefly, concentrated stock solutions of Z_R-$(ELS)_6$-C (10 mg/mL) were prepared by dissolving the lyophilized polymers in 8 M Urea, 100 mM NaH_2PO_4, 10 mM Tris, pH 7.2. Disulfide bonds between polymers were then reduced by adding 400 mM tris(2-carboxyethyl)phosphine (TCEP) and incubated at 37º C for 1.5 hours. Simultaneously, amine-labeled oligonucleotides, 20 bp in length, were reacted for one hour at 37º C with the heterobifunctional crosslinker, sulfosuccinimidyl 4-[*N*-maleimidomethyl]cyclohexane-1-carboxylate (sulfo-SMCC). Afterwards, the reduced artificial protein and the DNA solutions were passed through two separate NAP-5 desalting columns, combined, and incubated for 2 hours at room temperatures and then overnight at 4 C. The following day, the Z_R-$(ELS)_6$-C-ssDNA conjugates were isolated using a HiTrapQ ion-exchange column and fast protein liquid chromatography (FPLC, Figure 7.3). The peak fractions containing the conjugate were pooled and dialyzed with frequent water exchanges for 3 days and lyophilized. The polymers were then redissolved in Tris-EDTA (TE) buffer (pH 7.4) at a concentration of approximately 35 μM. The actual conjugate stock concentrations were determined from the OD_{260} using an extinction coefficient from the appended ssDNA's extinction coefficient (ε_{260} = 185,000 M^{-1} cm^{-1}). Conjugate stock solutions were aliquoted and stored at -20º C until needed.

7.2.5. DNA-Templated Protein Microarray Fabrication

DNA microarrays were fabricated by printing 5′-amine modified oligonucleotides on silyated (amine-aldehyde) functionalized slides (CEL Associates Inc). Oligonucleotide stocks were diluted appropriately with spotting buffer [40% (w/v) glycerol in PBS, pH 6.6]. For the single protein arrays, the arrays were printed in 90, 50, and 10 μM concentrations and a control strand was printed at 90 μM. The multi-protein arrays were all printed at 50 μM. A custom built microarrayer outfitted with SMP3 pins (Telechem Int. Sunnyvale, CA) was used for all printing. After printing, the slides were placed in a humid chamber for 45 minutes, and then allowed to dry outside the chamber for 10 minutes. The remaining free aldehyde groups were quenched in a solution of 25% (v/v) ethanol and 0.25% (w/v) $NaBH_4$ in PBS for 5 minutes. The slides were then thoroughly rinsed with milli-Q water (Millipore Billerica, MA) and blocked using pre-hybridization buffer [20 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM EDTA, 1% (v/v) Tween 20, 0.1 mg/mL denatured herring sperm DNA, 0.1 mg/mL BSA, 3% (w/v) dry milk] with gentle agitation for 2 hours. The arrays were then washed 3 times for 10 minutes each in wash buffer [20 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM EDTA, 0.01% (v/v) Tween 20], thoroughly rinsed with milli-Q water and dried under a nitrogen stream.

Artificial protein-DNA conjugates were used to direct the assembly of target proteins onto the DNA-spotted microarrays by first pre-incubating the target Z_E-fusion proteins individually with the artificial protein conjugates. A three-fold molar excess of target protein was typically maintained during these incubations. A reaction chamber was prepared by affixing a Gene Frame (AB Gene) to the microarrays. Mixtures of

artificial proteins and target proteins were brought to a final volume of 100 μL with hybridization buffer [20 mM Tris-HCl pH 7.4, 150 mM NaCl, 5 mM EDTA, 0.01% (v/v) Tween 20, 0.1 mg/mL denatured herring sperm DNA, 0.5% (w/v) dry milk] and perfused into the reaction chamber. For multi-protein array fabrication, two solutions of polymer-target protein complexes were prepared separately, combined into one solution, and introduced to the reaction chamber. The slides were left at room temperature to incubate overnight. Afterwards, the Gene Frames were removed and the slides were washed three times with wash buffer, rinsed with milli-Q water, and dried under nitrogen stream. New Gene Frames were applied to each array and a solution containing Alexa647-anti-GFP antibodies or both Alexa647-anti-GFP and Cy3-anti-GST antibodies were added at concentrations of 1 μg/mL for two hours. The slides were then washed 3 times for 10 minutes in wash buffer, rinsed briefly with milli-Q water and dried. The slides were immediately imaged on a GenePix 4000B microarray scanner (Molecular Devices Sunnyvale, CA) using a gain of 450.

The zipper exchange time studies were performed by pre-incubating the GFP-Z_E with a polymer-DNA conjugate and separately pre-incubating the GST-Z_E with an unconjugated polymer protein for 30 minutes. Following this, the two solutions were mixed together with hybridization buffer and incubated together for 1, 2, 4, or 8 hours. The solution was then incubated in a reaction chamber on a microarray printed with a control strand and a strand complementary to the GFP-Z_E/polymer-DNA complex and left overnight. The slides were then washed, probed, and scanned identically to the multi-protein microarrays.

7.2.6. Surface-Immobilized β-Lactamase Assays

All reagents used to prepare substrates for these assays were dissolved in TE buffer (pH 7.4). 96-well plates were incubated with biotin-BSA (1 mg/mL) for 10 minutes. The plates were then blocked with casein (1 mg/mL) followed by BSA (1 mg/mL) for 1 minute each, and washed twice with TE buffer. Next, streptavidin (0.5 mg/mL) was added to each well and incubated for 10 minutes. Afterwards, the streptavidin solutions were aspirated off and the plates were rinsed briefly with the casein solution and then twice with TE buffer. Biotin functionalized DNA was added to the appropriate wells to a final volume of 50 μL and incubated for 30 minutes. Certain wells were designated for controls that either omitted the DNA deposition or contained a non-complementary DNA strand. Excess DNA that did not bind the surface was removed by washing with casein and BSA as before. To immobilize enzymes, PSE-4-Z_E was pre-incubated in 3 fold excess of the Z_R-$(ELS)_6$-C-ssDNA conjugate for 30 minutes. The PSE-4-Z_E and conjugate solution were applied to the appropriate wells. The plate was left to incubate for 2 hours and then protein solutions were aspirated from the wells. After washing 4 times with TE buffer, each well was brought to a final volume of 100 μL by adding TE buffer. Subsequently, nitrocefin was added to each well, bringing the wells to a final volume of 200 μL and a final substrate concentration to 150 μM. The plates were scanned using a Tecan Infinite 200 microplate reader. Absorbance readings at 485 nm were monitored every 50 seconds for 3 hours at 25 C. The steady-state reaction rates of the immobilized β-Lactamase enzymes were determined via linear regression analysis.

Each condition was performed in triplicate (5 μg complementary DNA, 5 μg non-complementary DNA, and no DNA). The rates of nitrocefin hydrolysis were averaged

across the replicates and plotted with error bars indicating +/- standard deviations. Different amounts of DNA (0, 1, 1.5, 2, 4, or 5 μg) were added to wells in triplicate. The steady state rates were calculated via linear regression and calibrated against known concentrations of soluble PSE-4 to estimate the concentration of PSE-4 immobilized for each DNA concentration. The values were averaged and plotted with error bars indicating +/- standard deviations. These points were then fitted using linear regression. All statistical analyses were performed using OriginPro 8 (OriginLab).

7.2.7. Synthesis and Characterization of Multi-Protein Assemblies

The DNA scaffold consisted of a single duplex of DNA (34 bases; 10 nm in length) flanked by unique single-strand 'overhangs' comprised of 20 bases on either end. To create scaffolds, strands were mixed stoichiometrically as estimated by A_{260}, and annealed from 95º C to RT, over 1.5 hrs in TAE/Mg^{2+} buffer (40 mM Tris, 40 mM Acetic Acid, 2 mM EDTA, 12.5 mM Mg^{2+} Acetate, pH 8) for a final concentration of 1.8 μM. Target protein (either CaM-Z_E or GFP-Z_E) and its respective Z_R-$(ELS)_6$-C-ssDNA conjugate were mixed in a 1:1 ratio and incubated at 4º C for at least 1 hr. DNA scaffold was added in a 1:2 or 1:1 ratio with the protein mixture (depending on whether it was the single or two protein system) and incubated at RT for 1.5 hrs, then 4º C for 30 min. A 7% (w/v) non-denaturing PAGE gel was run at 200 V at 4º C (SE 600 Ruby, GE Healthcare) in TAE/Mg^{2+} buffer. The DNA/protein complexes were visualized by staining with Stains-All (Sigma).

7.3. Results and Discussion

In this report, we describe a new synthetic strategy to couple proteins to DNA-matrices that involves the production of DNA-conjugated polypeptide polymers **1** that function as capture probes by associating with recombinant affinity-tagged proteins in solution and then directing their deposition onto DNA-coated supports. These polymers are based on artificial proteins that have been previously used to control the surface immobilization of proteins [143] and to build finite-sized multi-protein complexes [144]. With this system, protein capture is achieved through the coiled-coil association of an engineered parallel pair of heterodimeric leucine zippers, designated Z_E and Z_R. The zipper sequences are derived from polypeptides developed by Vinson *et al.* [98], and form exceptionally strong heterodimeric complexes ($K_D \sim 10^{-15}$ M) with much weaker homodimeric complexes ($K_D \sim 10^{-3}$ to 10^{-6} M). Half of the zipper complex is fused to a target protein as an affinity tag, while the other component is incorporated into the polymer as a genetic fusion. The polymers also contain a mechanically flexible and repetitive domain based on the elastomeric poly(VPGV_αG) structural motif of the protein elastin (EL). Substitution of amino acids at the V_α position of this domain provides control over the hydrophilicity of the polymer, and can be used to either direct or minimize physical adhesion of the polymers to functionalized surfaces [143]. Here, we demonstrate that a DNA-conjugated version of these artificial proteins can be used to direct the self-assembly of target proteins onto DNA supports by forming a monovalent and stable intermediate linkage between immobilized ssDNA and the target proteins (Figure 7.2).

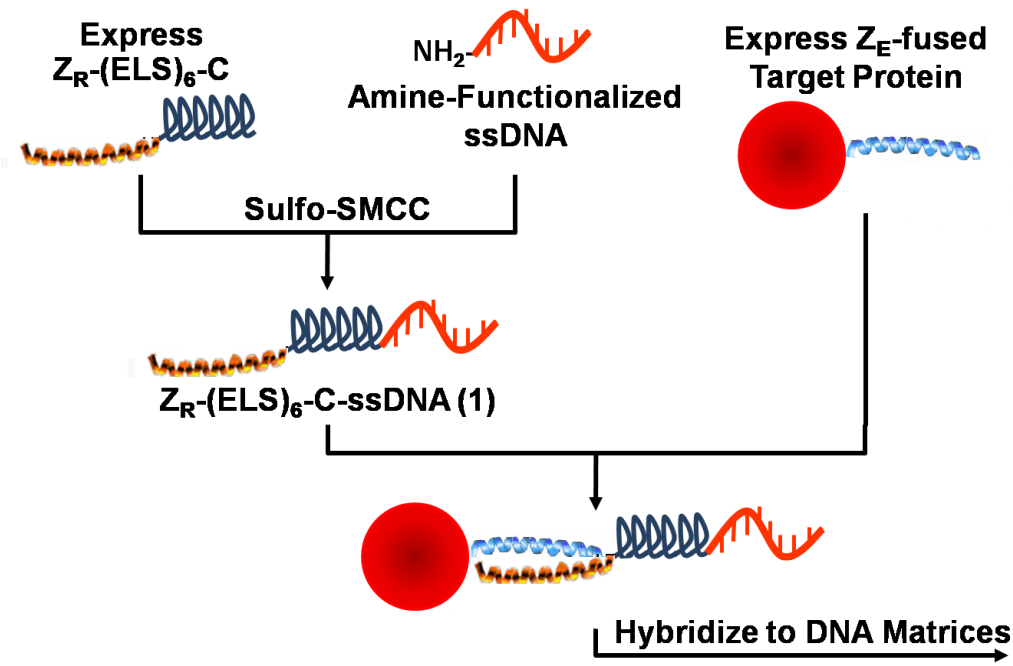

Figure 7.2. Synthetic Route to Employ Artificial Protein-DNA Conjugates for Recombinant Protein Capture

The polypeptide polymers employed here contain the basic portion of the zipper complex (Z_R), a polymerized elastic 'mid-block domain' or ELS fragment (VPGVG VPGSG VPGVG VPGSG VPGVG), and a C-terminal cysteine, yielding Z_R-$(ELS)_6$-C. After the *in vivo* expression and purification of these polymers, sulfosuccinimidyl 4-[*N*-maleimidomethyl]cyclohexane-1-carboxylate (sulfo-SMCC) was used as a heterobifunctional crosslinker to couple an amine-functionalized oligonucleotide to the C-terminal cysteine of the polymers [134]. The large shift in the isoelectric point of the polymers after DNA-labeling allowed the conjugates to be purified to homogeneity using fast protein liquid chromatography (Figure 7.3). Since the artificial proteins can sustain, without loss of function, lyophilization, resuspension in denaturing buffers, and extended dialysis, the labeling reactions can be scaled-up, and this route allows concentrated stock

solutions of Z_R-$(ELS)_6$-C-ssDNA to be prepared despite the intrinsic inefficiency of the sulfo-SMCC crosslinking reaction.

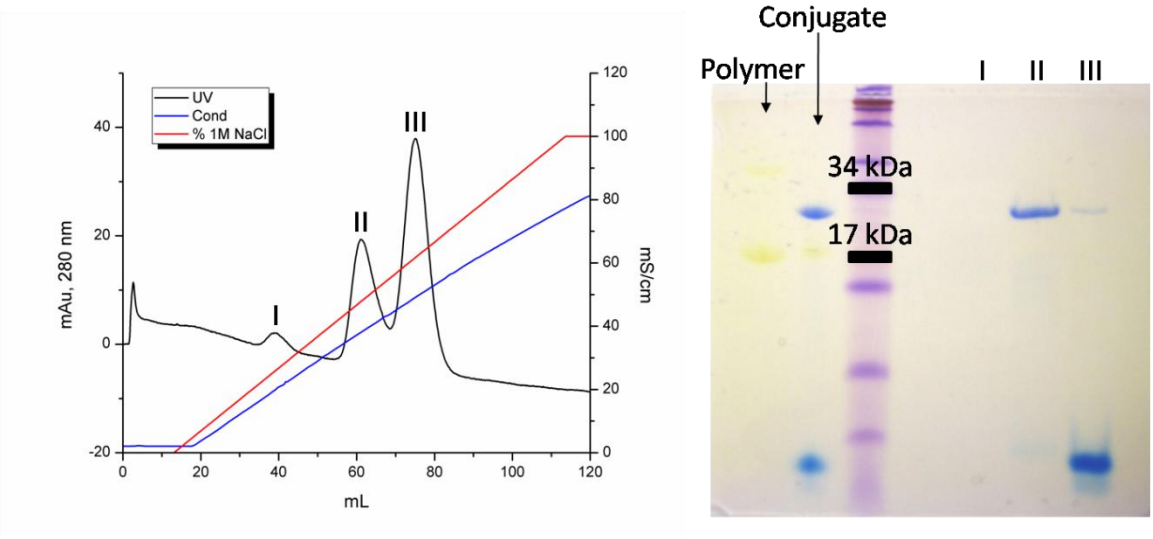

Figure 7.3 FPLC chromatogram and SDS-PAGE gel for conjugate synthesis and purification. The FPLC trace indicates the absorption measured at 280 nm, gradient implemented, and measured conductivity. The chromatogram shows the characteristic 3 peaks from the crosslinking reaction: I. Unconjugated polymer; II. Conjugate; III. Unconjugated DNA. The corresponding gel shows purified polymer, the crosslinking reaction products (Conjugate), and a sample of each of the 3 peaks. The Stains-All stains DNA blue and protein yellow. The shift in the DNA band shows the presence of the conjugate in peak II.

To test the ability of **1** to control the deposition of protein on DNA matrices, we labeled two samples of Z_R-$(ELS)_6$-C with different oligonucleotides containing non-complementary sequences. These polymers were then used as solution-phase capture probes that encode the spatial deposition of recombinant proteins onto DNA-printed microarrays (Figure 7.4). In this procedure, green fluorescent protein (GFP) and glutathione-*S*-transferase (GST) were chosen as target proteins and were expressed *in*

vivo with C-terminal Z_E fusions. Protein capture was achieved by mixing samples of either Ni^{2+}-NTA purified target proteins (~1 mg/mL) or expressed proteins in unpurified cell lysates with a single conjugated version of **1** at a 3:1 molar ratio of target protein to polymer. After a brief incubation period (< 30 min), these solutions were incubated overnight over a custom fabricated DNA microarray. Each array was washed and probed with fluorescently labeled antibodies (either Alexa647-anti-GFP or both Alexa647-anti-GFP and Cy3-anti-GST at 1 μg/mL) and scanned using a GenePix 4000B microarray scanner (Molecular Devices). DNA-directed self-assembly of both GFP-Z_E and GST-Z_E was found to be highly selective. As expected, arrays formed using either purified proteins or cell lysates demonstrated the oligonucleotide sequence-specific targeting of the proteins. Microarray spots containing the appropriate complementary ssDNA sequence possessed average signal-to-noise ratios (SNR) with a range of 20 to 60. Proteins appearing in the Cy3 channel tended towards the higher end of this range. On the other hand, average SNR values were significantly lower for spots containing non-complementary strands, a SNR range of 1.6 to 5. Furthermore, when the pre-incubated protein solutions were combined for 1, 2, 4, or 8 hours prior to the overnight incubation on the microarray, we did not observe any detectable exchange between proteins displaying the zipper fusion (Figure 7.5).

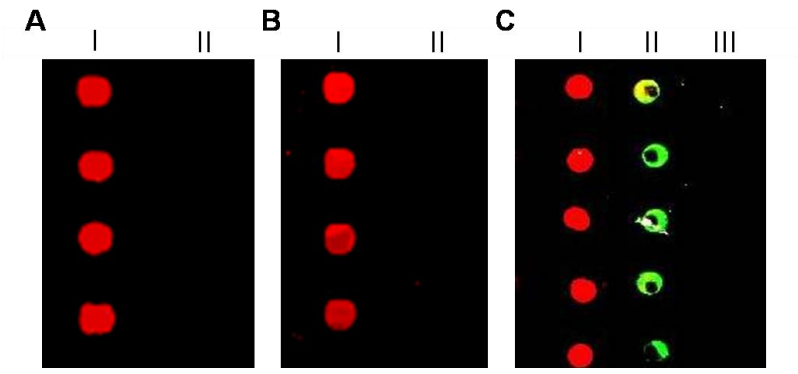

Figure 7.4 The use of artificial polypeptide conjugates to create protein microarrays encoded by DNA. Arrays probed with Alexa647-anti-GFP were prepared using purified GFP-Z_E (A) and cell lysates containing expressed GFP-Z_E (B). In both cases, column I denotes spots containing the complementary sequence for GFP-Z_E / polymer complex while a non-complementary control strand was spotted in column II. Images in A and B were cropped and placed together for clarity with a pitch of 250 μm between spots. (C) Two-component arrays were fabricated through a single incubation of pre-captured GFP-Z_E and GST-Z_E proteins. Column I contains a complementary strand for GFP-Z_E / polymer complex, column II contains a complementary sequence for a second GST-Z_E / polymer complex, and III contains spots of a non-complementary control strand. Multi-protein arrays were probed with a mixture of Alexa647-anti-GFP and Cy3-anti-GST. The pitch between spots is 500 μm.

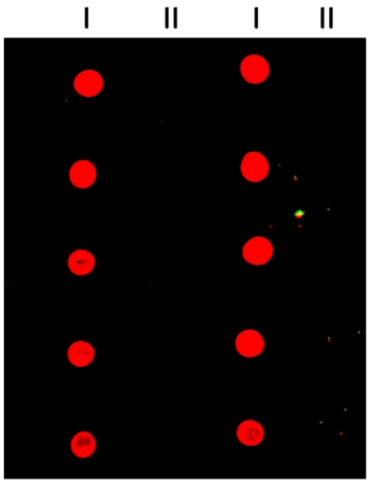

Figure 7.5 Microarray experiment showing the leucine zippers do not exchange between target proteins. Two protein complexes: (1) a GFP-Z_E protein captured using a DNA-polypeptide conjugate, and (2) a GST-Z_E complexed with an unconjugated Z_R-$(ELS)_6$-C polypeptide were mixed together for periods of 1-8 hours to test exchange. Afterwards, this solution was incubated for 16 hours over DNA-spotted array. Without exchange, the GST-Z_E protein should not immobilize on our arrays, and any exchange should result in the appearance of a GST positive signal. All of the arrays show selective capture of the GFP-Z_E. We did not observe evidence of exchange in any experiments. The 8 hour time point is shown in the figure. The positive control strand was printed in lane I (the GFP positive lane), while a non-complementary strand is was printed in lane II: The GST signal (green) was lower than background in all spots in the array.

We also examined the effect our DNA-directed immobilization strategy had on enzymatic activity. As a test case, we immobilized a β-lactamase enzyme (PSE-4) onto the surface of 96-well plates that were functionalized with streptavidin and then biotin-labeled oligonucleotides. DNA-directed deposition was performed using a similar procedure to that developed for the microarrays except that the PSE-4-Z_E fusion protein was prepared using standard periplasmic expression and lysis methods. Nitrocefin was used as a substrate for PSE-4 since the enzymatic opening of its β–lactam bond produced

a detectable change in absorption at 485 nm using a Tecan Infinite 200 microplate reader. Measurements of nitrocefin hydrolysis rates again verified ability of **1** to selectively control enzyme deposition of target proteins (Figure 7.6B). Wells containing 5 μg of a complementary oligonucleotide to **1** showed a 4-fold increase in activity over wells that contained non-complementary sequences or where the ssDNA was omitted (Figure 7.6). Consistent with previous reports that utilize DNA-directed immobilization of enzymes [132], levels of PSE-4 activity were found to be linearly dependent on the amount of DNA that is immobilized in the wells (Figure 7.6 inset). Using soluble PSE-4-Z_E as a calibration standard, the hydrolysis rate was converted to PSE-4-Z_E concentration, yielding a slope of 18.6 nM PSE-4-Z_E per microgram of DNA.

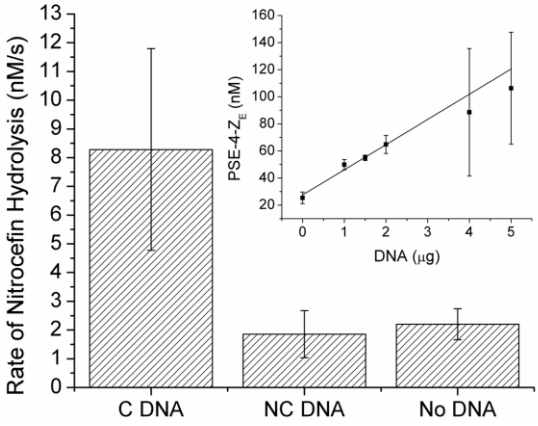

Figure 7.6 DNA-directed control over immobilized enzyme activity. A 4-fold increase in PSE-4-Z_E activity over background nitrocefin hydrolysis is seen in wells that were incubated with 5 μg of complementary DNA strands (C DNA) when compared to wells containing the same amount of a non-complementary DNA strand (NC DNA). Control wells where the DNA-functionalization step was omitted (No DNA) are also shown for comparison. The inset shows the linear dependence of PSE-4-Z_E activity on the amount of complementary DNA deposited in the wells (r^2 = 0.95).

Finally, the two different conjugated versions of **1** were synthesized and used to facilitate the self-assembly of multi-protein complexes onto molecular-scale scaffolds formed from DNA. Complexes containing either two GFP-Z_E proteins or a GFP-Z_E and a calmodulin protein (CaM-Z_E) were prepared (Figure 7.7). The DNA scaffold consisted of a 10 nm long double helix (34 bases) flanked by unique single-strand 'overhangs' comprised of 20 bases on either end. The 'overhangs' either possessed the same or orthogonal oligonucleotide sequences depending on whether single or multi-component complexes were being synthesized. Multi-protein assemblies were synthesized by pre-incubating **1** with GFP-Z_E or CaM-Z_E, and then adding the DNA scaffold. In each case, the protein, polymer, and 'overhang' concentrations were held at a 1:1:1 stoichiometry. Complex formation was examined using non-denaturing polyacrylamide gel electorphoresis (PAGE) and by staining for DNA and protein using Stains-All (Sigma). Clear separation between the scaffold and both partially and fully-formed assemblies was observed. For each complex, gel intensity analysis indicated that the multi-protein complexes formed in high yield (> 95%) using these conditions, proving that the association of the zipper complex and DNA-hybridization provided stable and selective linkages to drive the assembly process near quantitative yield, and that this process does not require the addition of large excesses of individual assembly components to form complete complexes.

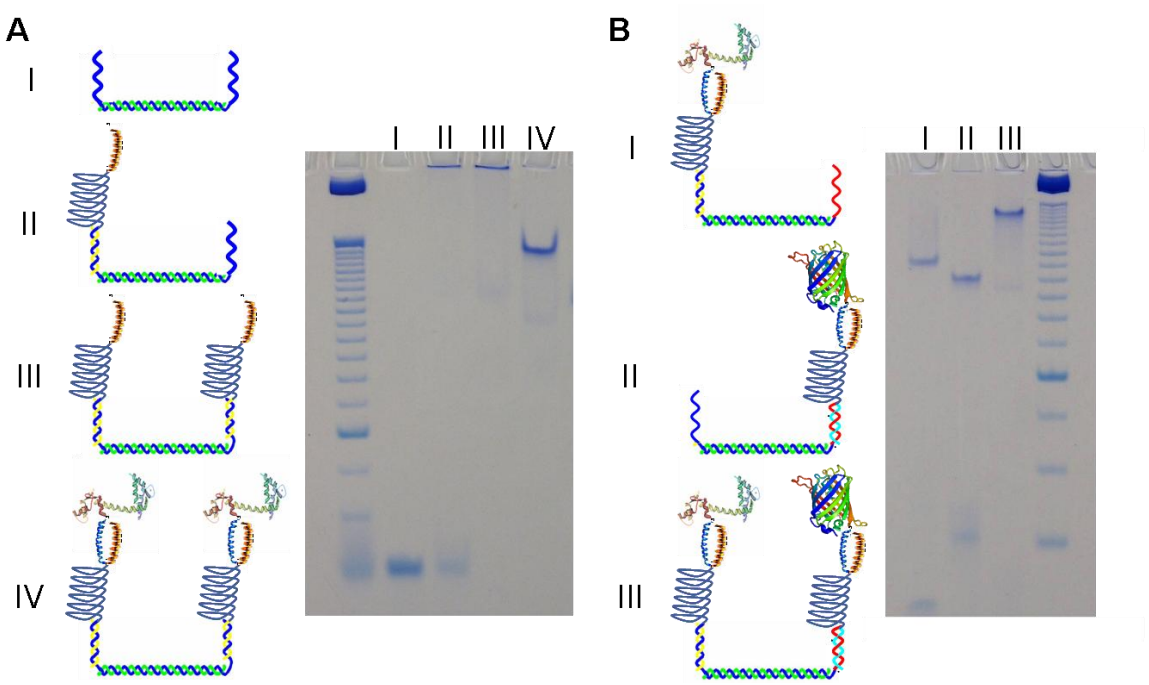

Figure 7.7 Non-denaturing PAGE-gels confirming the synthesis of multiprotein complexes. (A) Native PAGE-gel showing the successful formation of the complex containing two Cam-Z_E proteins assembled on a single DNA double helix. The dominant species observed in the individual lanes of the gel are indicated by Roman numerals. The labeling scheme is as follows: I. DNA scaffold alone; II. assemblies prepared at a 1 to 0.5 ratio of ssDNA scaffold to conjugates; III. a 'half' complex composed of the DNA scaffolds and two artificial polypeptides; IV. complete assemblies formed at a 1:1:1 ratio of all assembly components. (B) Gel demonstrating the assembly of multi-protein complexes containing both a GFP-Z_E and a Cam-Z_E (III), templated on a DNA scaffold possessing two different 'overhang' sequences. Partial complexes formed using only one type of protein and corresponding to either Cam-Z_E (I), or GFP-Z_E (II) are also shown. A full version of the gels and complexes seen here can be found in Appendix A.

7.4. Conclusions

In summary, we have demonstrated that DNA-conjugated artificial polypeptides can be used as protein capture probes that facilitate the oligonucleotide sequence-dependent deposition onto DNA-functionalized matrices. Furthermore, we have shown

that these polymers allow the stoichiometries of enzymes to be deterministically controlled by forming a stable intermediate linkage between DNA-coated supports and target proteins. Importantly, our modular strategy separates the production of recombinant target proteins from DNA-conjugation procedures. Here, the synthesis of a single set of polymer allows diverse combinations of Z_E-tagged proteins to be immobilized for different applications. Furthermore, the selectivity and stability of the heterodimeric association of the leucine zipper system employed here permits target protein capture and pull down from unpurified cellular lysates and is compatible with a variety of protein expression systems and techniques. In this way, our strategy presents a flexible and versatile alternative to existing DNA-conjugation methods that is advantageous for applications where the ability to conjugate proteins with DNA is limiting, and in circumstances that require minimal post-expression processing of target proteins.

References

1. Arora, P. S.; Yamagiwa, H.; Srivastava, A.; Bolander, M. E.; Sarkar, G., Comparative evaluation of two two-dimensional gel electrophoresis image analysis software applications using synovial fluids from patients with joint disease. *J Orthop Sci* **2005,** *10* (2), 160-6.

2. Aulak, K. S.; Koeck, T.; Crabb, J. W.; Stuehr, D. J., Proteomic method for identification of tyrosine-nitrated proteins. *Methods Mol Biol* **2004,** *279*, 151-65.

3. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422* (6928), 198-207.

4. Kusnezow, W.; Banzon, V.; Schroder, C.; Schaal, R.; Hoheisel, J. D.; Ruffer, S.; Luft, P.; Duschl, A.; Syagailo, Y. V., Antibody microarray-based profiling of complex specimens: systematic evaluation of labeling strategies. *Proteomics* **2007,** *7* (11), 1786-99.

5. Lequin, R. M., Enzyme immunoassay (EIA)/enzyme-linked immunosorbent assay (ELISA). *Clin Chem* **2005,** *51* (12), 2415-8.

6. Anderson, A. R.; Weaver, A. M.; Cummings, P. T.; Quaranta, V., Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. *Cell* **2006,** *127* (5), 905-15.

7. Ford, B. K.; Volin, C. E.; Murphy, S. M.; Lynch, R. M.; Descour, M. R., Computed tomography-based spectral imaging for fluorescence microscopy. *Biophys J* **2001,** *80* (2), 986-93.

8. Gao, L.; Kester, R. T.; Hagen, N.; Tkaczyk, T. S., Snapshot Image Mapping Spectrometer (IMS) with high sampling density for hyperspectral microscopy. *Opt Express* **2010,** *18* (14), 14330-44.

9. Hiraoka, Y.; Shimi, T.; Haraguchi, T., Multispectral imaging fluorescence microscopy for living cells. *Cell Struct Funct* **2002,** *27* (5), 367-74.

10. Hutter, H., Five-colour in vivo imaging of neurons in Caenorhabditis elegans. *J Microsc* **2004,** *215* (Pt 2), 213-8.

11. Tsurui, H.; Nishimura, H.; Hattori, S.; Hirose, S.; Okumura, K.; Shirai, T., Seven-color fluorescence imaging of tissue samples based on Fourier spectroscopy and singular value decomposition. *J Histochem Cytochem* **2000,** *48* (5), 653-62.

12. Alaiya, A. A.; Franzen, B.; Auer, G.; Linder, S., Cancer proteomics: from identification of novel markers to creation of artifical learning models for tumor classification. *Electrophoresis* **2000,** *21* (6), 1210-7.

13. Heppner, G. H.; Miller, B. E., Tumor heterogeneity: biological implications and therapeutic consequences. *Cancer Metastasis Rev* **1983,** *2* (1), 5-23.

14. Liu, J.; Lau, S. K.; Varma, V. A.; Moffitt, R. A.; Caldwell, M.; Liu, T.; Young, A. N.; Petros, J. A.; Osunkoya, A. O.; Krogstad, T.; Leyland-Jones, B.; Wang, M. D.; Nie, S., Molecular mapping of tumor heterogeneity on clinical tissue specimens with multiplexed quantum dots. *ACS Nano* **2010,** *4* (5), 2755-65.

15. Cregger, M.; Berger, A. J.; Rimm, D. L., Immunohistochemistry and quantitative analysis of protein expression. *Arch Pathol Lab Med* **2006,** *130* (7), 1026-30.

16. Bast, R. C., Jr.; Ravdin, P.; Hayes, D. F.; Bates, S.; Fritsche, H., Jr.; Jessup, J. M.; Kemeny, N.; Locker, G. Y.; Mennel, R. G.; Somerfield, M. R., 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines of the American Society of Clinical Oncology. *J Clin Oncol* **2001,** *19* (6), 1865-78.

17. Hayes, D. F.; Bast, R. C.; Desch, C. E.; Fritsche, H., Jr.; Kemeny, N. E.; Jessup, J. M.; Locker, G. Y.; Macdonald, J. S.; Mennel, R. G.; Norton, L.; Ravdin, P.; Taube, S.; Winn, R. J., Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst* **1996,** *88* (20), 1456-66.

18. Schilsky, R. L.; Taube, S. E., Tumor markers as clinical cancer tests--are we there yet? *Semin Oncol* **2002,** *29* (3), 211-2.

19. Duose, D. Y. Multiplexed and Reiterative Detection of Protein Markers in Cells using Dynamic Nucleic Acid Complexes. Rice University, Houston, TX, 2011.

20. Klonisch, T.; Wiechec, E.; Hombach-Klonisch, S.; Ande, S. R.; Wesselborg, S.; Schulze-Osthoff, K.; Los, M., Cancer stem cell markers in common cancers - therapeutic implications. *Trends Mol Med* **2008,** *14* (10), 450-60.

21. Woodward, W. A.; Sulman, E. P., Cancer stem cells: markers or biomarkers? *Cancer Metastasis Rev* **2008,** *27* (3), 459-70.

22. Al-Hajj, M.; Wicha, M. S.; Benito-Hernandez, A.; Morrison, S. J.; Clarke, M. F., Prospective identification of tumorigenic breast cancer cells. *Proc Natl Acad Sci U S A* **2003,** *100* (7), 3983-8.

23. Burger, P. E.; Xiong, X.; Coetzee, S.; Salm, S. N.; Moscatelli, D.; Goto, K.; Wilson, E. L., Sca-1 expression identifies stem cells in the proximal region of prostatic

ducts with high capacity to reconstitute prostatic tissue. *Proc Natl Acad Sci U S A* **2005,** *102* (20), 7180-5.

24. Ouhtit, A.; Abd Elmageed, Z. Y.; Abdraboh, M. E.; Lioe, T. F.; Raj, M. H., In vivo evidence for the role of CD44s in promoting breast cancer metastasis to the liver. *Am J Pathol* **2007,** *171* (6), 2033-9.

25. Patrawala, L.; Calhoun, T.; Schneider-Broussard, R.; Li, H.; Bhatia, B.; Tang, S.; Reilly, J. G.; Chandra, D.; Zhou, J.; Claypool, K.; Coghlan, L.; Tang, D. G., Highly purified CD44+ prostate cancer cells from xenograft human tumors are enriched in tumorigenic and metastatic progenitor cells. *Oncogene* **2006,** *25* (12), 1696-708.

26. Roskams, T., Liver stem cells and their implication in hepatocellular and cholangiocarcinoma. *Oncogene* **2006,** *25* (27), 3818-22.

27. Shackleton, M.; Vaillant, F.; Simpson, K. J.; Stingl, J.; Smyth, G. K.; Asselin-Labat, M. L.; Wu, L.; Lindeman, G. J.; Visvader, J. E., Generation of a functional mammary gland from a single stem cell. *Nature* **2006,** *439* (7072), 84-8.

28. Signoretti, S.; Waltregny, D.; Dilks, J.; Isaac, B.; Lin, D.; Garraway, L.; Yang, A.; Montironi, R.; McKeon, F.; Loda, M., p63 is a prostate basal cell marker and is required for prostate development. *Am J Pathol* **2000,** *157* (6), 1769-75.

29. Bao, S.; Wu, Q.; McLendon, R. E.; Hao, Y.; Shi, Q.; Hjelmeland, A. B.; Dewhirst, M. W.; Bigner, D. D.; Rich, J. N., Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **2006,** *444* (7120), 756-60.

30. Phillips, T. M.; McBride, W. H.; Pajonk, F., The response of CD24(-/low)/CD44+ breast cancer-initiating cells to radiation. *J Natl Cancer Inst* **2006,** *98* (24), 1777-85.

31. Bobrow, M. N.; Harris, T. D.; Shaughnessy, K. J.; Litt, G. J., Catalyzed reporter deposition, a novel method of signal amplification. Application to immunoassays. *J Immunol Methods* **1989,** *125* (1-2), 279-85.

32. Gao, X.; Nie, S., Molecular profiling of single cells and tissue specimens with quantum dots. *Trends Biotechnol* **2003,** *21* (9), 371-3.

33. Pirici, D.; Mogoanta, L.; Kumar-Singh, S.; Pirici, I.; Margaritescu, C.; Simionescu, C.; Stanescu, R., Antibody elution method for multiple immunohistochemistry on primary antibodies raised in the same species and of the same subtype. *J Histochem Cytochem* **2009,** *57* (6), 567-75.

34. Toth, Z. E.; Mezey, E., Simultaneous visualization of multiple antigens with tyramide signal amplification using antibodies from the same species. *J Histochem Cytochem* **2007,** *55* (6), 545-54.

35. Femino, A. M.; Fay, F. S.; Fogarty, K.; Singer, R. H., Visualization of single RNA transcripts in situ. *Science* **1998,** *280* (5363), 585-90.

36. Itzkovitz, S.; Lyubimova, A.; Blat, I. C.; Maynard, M.; van Es, J.; Lees, J.; Jacks, T.; Clevers, H.; van Oudenaarden, A., Single-molecule transcript counting of stem-cell markers in the mouse intestine. *Nat Cell Biol* **2012,** *14* (1), 106-14.

37. Lawrence, J. B.; Singer, R. H.; Marselle, L. M., Highly localized tracks of specific transcripts within interphase nuclei visualized by in situ hybridization. *Cell* **1989,** *57* (3), 493-502.

38. Pinkel, D.; Straume, T.; Gray, J. W., Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proc Natl Acad Sci U S A* **1986,** *83* (9), 2934-8.

39. Sundell, C. L.; Singer, R. H., Requirement of microfilaments in sorting of actin messenger RNA. *Science* **1991,** *253* (5025), 1275-7.

40. Zhang, G.; Taneja, K. L.; Singer, R. H.; Green, M. R., Localization of pre-mRNA splicing in mammalian nuclei. *Nature* **1994,** *372* (6508), 809-12.

41. Fusco, D.; Accornero, N.; Lavoie, B.; Shenoy, S. M.; Blanchard, J. M.; Singer, R. H.; Bertrand, E., Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr Biol* **2003,** *13* (2), 161-7.

42. Shav-Tal, Y.; Darzacq, X.; Shenoy, S. M.; Fusco, D.; Janicki, S. M.; Spector, D. L.; Singer, R. H., Dynamics of single mRNPs in nuclei of living cells. *Science* **2004,** *304* (5678), 1797-800.

43. Park, H. Y.; Buxbaum, A. R.; Singer, R. H., Single mRNA tracking in live cells. *Methods Enzymol* **2010,** *472*, 387-406.

44. Bratu, D. P.; Cha, B. J.; Mhlanga, M. M.; Kramer, F. R.; Tyagi, S., Visualizing the distribution and transport of mRNAs in living cells. *Proc Natl Acad Sci U S A* **2003,** *100* (23), 13308-13.

45. Santangelo, P. J.; Nix, B.; Tsourkas, A.; Bao, G., Dual FRET molecular beacons for mRNA detection in living cells. *Nucleic Acids Res* **2004,** *32* (6), e57.

46. Sokol, D. L.; Zhang, X.; Lu, P.; Gewirtz, A. M., Real time detection of DNA.RNA hybridization in living cells. *Proc Natl Acad Sci U S A* **1998,** *95* (20), 11538-43.

47. Tsuji, A.; Sato, Y.; Hirano, M.; Suga, T.; Koshimoto, H.; Taguchi, T.; Ohsuka, S., Development of a time-resolved fluorometric method for observing hybridization in living cells using fluorescence resonance energy transfer. *Biophys J* **2001,** *81* (1), 501-15.

48. Tyagi, S.; Kramer, F. R., Molecular beacons: probes that fluoresce upon hybridization. *Nat Biotechnol* **1996,** *14* (3), 303-8.

49. Kang, W. J.; Chae, J. R.; Cho, Y. L.; Lee, J. D.; Kim, S., Multiplex imaging of single tumor cells using quantum-dot-conjugated aptamers. *Small* **2009,** *5* (22), 2519-22.

50. Shangguan, D.; Li, Y.; Tang, Z.; Cao, Z. C.; Chen, H. W.; Mallikaratchy, P.; Sefah, K.; Yang, C. J.; Tan, W., Aptamers evolved from live cells as effective molecular probes for cancer study. *Proc Natl Acad Sci U S A* **2006,** *103* (32), 11838-43.

51. Stanlis, K. K.; McIntosh, J. R., Single-strand DNA aptamers as probes for protein localization in cells. *J Histochem Cytochem* **2003,** *51* (6), 797-808.

52. Sano, T.; Smith, C. L.; Cantor, C. R., Immuno-PCR: very sensitive antigen detection by means of specific antibody-DNA conjugates. *Science* **1992,** *258* (5079), 120-2.

53. Kazane, S. A.; Sok, D.; Cho, E. H.; Uson, M. L.; Kuhn, P.; Schultz, P. G.; Smider, V. V., Site-specific DNA-antibody conjugates for specific and sensitive immuno-PCR. *Proc Natl Acad Sci U S A* **2012,** *109* (10), 3731-6.

54. Schweitzer, B.; Wiltshire, S.; Lambert, J.; O'Malley, S.; Kukanskis, K.; Zhu, Z.; Kingsmore, S. F.; Lizardi, P. M.; Ward, D. C., Immunoassays with rolling circle DNA amplification: a versatile platform for ultrasensitive antigen detection. *Proc Natl Acad Sci U S A* **2000,** *97* (18), 10113-9.

55. Fredriksson, S.; Dixon, W.; Ji, H.; Koong, A. C.; Mindrinos, M.; Davis, R. W., Multiplexed protein detection by proximity ligation for cancer biomarker validation. *Nat Methods* **2007,** *4* (4), 327-9.

56. Fredriksson, S.; Gullberg, M.; Jarvius, J.; Olsson, C.; Pietras, K.; Gustafsdottir, S. M.; Ostman, A.; Landegren, U., Protein detection using proximity-dependent DNA ligation assays. *Nat Biotechnol* **2002,** *20* (5), 473-7.

57. Teo, Y. N.; Wilson, J. N.; Kool, E. T., Polyfluorophores on a DNA backbone: a multicolor set of labels excited at one wavelength. *J Am Chem Soc* **2009,** *131* (11), 3923-33.

58. Guo, J.; Wang, S.; Dai, N.; Teo, Y. N.; Kool, E. T., Multispectral labeling of antibodies with polyfluorophores on a DNA backbone and application in cellular imaging. *Proc Natl Acad Sci U S A* **2011,** *108* (9), 3493-8.

59. Macdonald, J.; Li, Y.; Sutovic, M.; Lederman, H.; Pendri, K.; Lu, W.; Andrews, B. L.; Stefanovic, D.; Stojanovic, M. N., Medium scale integration of molecular logic gates in an automaton. *Nano Lett* **2006,** *6* (11), 2598-603.

60. Seelig, G.; Soloveichik, D.; Zhang, D. Y.; Winfree, E., Enzyme-free nucleic acid logic circuits. *Science* **2006,** *314* (5805), 1585-8.

61. Green, S. J.; Bath, J.; Turberfield, A. J., Coordinated chemomechanical cycles: a mechanism for autonomous molecular motion. *Phys Rev Lett* **2008,** *101* (23), 238101.

62. Venkataraman, S.; Dirks, R. M.; Rothemund, P. W.; Winfree, E.; Pierce, N. A., An autonomous polymerization motor powered by DNA hybridization. *Nat Nanotechnol* **2007,** *2* (8), 490-4.

63. Yin, P.; Choi, H. M. T.; Calvert, C. R.; Pierce, N. A., Programming biomolecular self-assembly pathways. *Nature* **2008,** *451* (7176), 318-22.

64. Yurke, B.; Turberfield, A. J.; Mills, A. P.; Simmel, F. C.; Neumann, J. L., A DNA-fuelled molecular machine made of DNA. *Nature* **2000,** *406* (6796), 605-8.

65. Hagiya, M.; Yaegashi, S.; Takahashi, K., Computing with Hairpins and Secondary Structures of DNA. In *Nanotechnology: Science and Computation*, Chen, J.; Jonoska, N.; Rozenberg, G., Eds. Springer: 2006; pp 293-308.

66. Qian, L.; Winfree, E., Scaling up digital circuit computation with DNA strand displacement cascades. *Science* **2011,** *332* (6034), 1196-201.

67. Qian, L.; Winfree, E.; Bruck, J., Neural network computation with DNA strand displacement cascades. *Nature* **2011,** *475* (7356), 368-72.

68. Zhang, D. Y.; Turberfield, A. J.; Yurke, B.; Winfree, E., Engineering entropy-driven reactions and networks catalyzed by DNA. *Science* **2007,** *318* (5853), 1121-5.

69. Seferos, D. S.; Giljohann, D. A.; Hill, H. D.; Prigodich, A. E.; Mirkin, C. A., Nano-flares: probes for transfection and mRNA detection in living cells. *J Am Chem Soc* **2007,** *129* (50), 15477-9.

70. Prigodich, A. E.; Randeria, P. S.; Briley, W. E.; Kim, N. J.; Daniel, W. L.; Giljohann, D. A.; Mirkin, C. A., Multiplexed nanoflares: mRNA detection in live cells. *Anal Chem* **2012,** *84* (4), 2062-6.

71. Dirks, R. M.; Pierce, N. A., Triggered amplification by hybridization chain reaction. *Proc Natl Acad Sci U S A* **2004,** *101* (43), 15275-8.

72. Choi, H. M. T.; Chang, J. Y.; Trinh, L. A.; Padilla, J. E.; Fraser, S. E.; Pierce, N. A., Programmable in situ amplification for multiplexed imaging of mRNA expression. *Nat Biotechnol* **2010,** *28* (11), 1208-12.

73. Schubert, W.; Bonnekoh, B.; Pommer, A. J.; Philipsen, L.; Bockelmann, R.; Malykh, Y.; Gollnick, H.; Friedenberger, M.; Bode, M.; Dress, A. W., Analyzing

proteome topology and function by automated multidimensional fluorescence microscopy. *Nat Biotechnol* **2006,** *24* (10), 1270-8.

74. Suzuki, T.; Tate, G.; Ikeda, K.; Mitsuya, T., A novel multicolor immunofluorescence method using heat treatment. *Acta Med Okayama* **2005,** *59* (4), 145-51.

75. Mitchell, P., Turning the spotlight on cellular imaging. *Nat Biotechnol* **2001,** *19* (11), 1013-7.

76. Nakakuki, T.; Birtwistle, M. R.; Saeki, Y.; Yumoto, N.; Ide, K.; Nagashima, T.; Brusch, L.; Ogunnaike, B. A.; Okada-Hatakeyama, M.; Kholodenko, B. N., Ligand-specific c-Fos expression emerges from the spatiotemporal control of ErbB network dynamics. *Cell* **2010,** *141* (5), 884-96.

77. Wu, X.; Liu, H.; Liu, J.; Haley, K. N.; Treadway, J. A.; Larson, J. P.; Ge, N.; Peale, F.; Bruchez, M. P., Immunofluorescent labeling of cancer marker Her2 and other cellular targets with semiconductor quantum dots. *Nat Biotechnol* **2003,** *21* (1), 41-6.

78. Rhee, W. J.; Santangelo, P. J.; Jo, H.; Bao, G., Target accessibility and signal specificity in live-cell detection of BMP-4 mRNA using molecular beacons. *Nucleic Acids Res* **2008,** *36* (5), e30.

79. Soderberg, O.; Gullberg, M.; Jarvius, M.; Ridderstrale, K.; Leuchowius, K. J.; Jarvius, J.; Wester, K.; Hydbring, P.; Bahram, F.; Larsson, L. G.; Landegren, U., Direct observation of individual endogenous protein complexes in situ by proximity ligation. *Nat Methods* **2006,** *3* (12), 995-1000.

80. Hoos, A.; Stojadinovic, A.; Singh, B.; Dudas, M. E.; Leung, D. H.; Shaha, A. R.; Shah, J. P.; Brennan, M. F.; Cordon-Cardo, C.; Ghossein, R., Clinical significance of molecular expression profiles of Hurthle cell tumors of the thyroid gland analyzed via tissue microarrays. *Am J Pathol* **2002,** *160* (1), 175-83.

81. Roederer, M.; Brenchley, J. M.; Betts, M. R.; De Rosa, S. C., Flow cytometric analysis of vaccine responses: how many colors are enough? *Clin Immunol* **2004,** *110* (3), 199-205.

82. Turberfield, A. J.; Mitchell, J. C.; Yurke, B.; Mills, A. P., Jr.; Blakey, M. I.; Simmel, F. C., DNA fuel for free-running nanomachines. *Phys Rev Lett* **2003,** *90* (11), 118102.

83. Yurke, B.; Mills, A. P., Jr.; Cheng, S. L., DNA implementation of addition in which the input strands are separate from the operator strands. *Biosystems* **1999,** *52* (1-3), 165-74.

84. Schweller, R. M.; Constantinou, P. E.; Frankel, N. W.; Narayan, P.; Diehl, M. R., Design of DNA-conjugated polypeptide-based capture probes for the anchoring of proteins to DNA matrices. *Bioconjug Chem* **2008,** *19* (12), 2304-7.

85. Zhang, D. Y.; Winfree, E., Control of DNA strand displacement kinetics using toehold exchange. *J Am Chem Soc* **2009,** *131* (47), 17303-14.

86. Carlon, E.; Heim, T., Thermodynamics of RNA/DNA hybridization in high-density oligonucleotide microarrays. *Physica a-Statistical Mechanics and Its Applications* **2006,** *362* (2), 433-449.

87. Peterson, A. W.; Wolf, L. K.; Georgiadis, R. M., Hybridization of mismatched or partially matched DNA at surfaces. *J Am Chem Soc* **2002,** *124* (49), 14601-7.

88. Zhang, D. Y.; Winfree, E., Robustness and modularity properties of a non-covalent DNA catalytic reaction. *Nucleic Acids Res* **2010,** *38* (12), 4182-97.

89. Stojanovic, M. N.; Mitchell, T. E.; Stefanovic, D., Deoxyribozyme-based logic gates. *J Am Chem Soc* **2002,** *124* (14), 3555-61.

90. Lubrich, D.; Green, S. J.; Turberfield, A. J., Kinetically controlled self-assembly of DNA oligomers. *J Am Chem Soc* **2009,** *131* (7), 2422-3.

91. Goodman, R. P.; Heilemann, M.; Doose, S.; Erben, C. M.; Kapanidis, A. N.; Turberfield, A. J., Reconfigurable, braced, three-dimensional DNA nanostructures. *Nat Nanotechnol* **2008,** *3* (2), 93-6.

92. Zhang, D. Y.; Seelig, G., Dynamic DNA nanotechnology using strand-displacement reactions. *Nat Chem* **2011,** *3* (2), 103-13.

93. Zhang, D. Y., Cooperative hybridization of oligonucleotides. *J Am Chem Soc* **2011,** *133* (4), 1077-86.

94. Benenson, Y.; Gil, B.; Ben-Dor, U.; Adar, R.; Shapiro, E., An autonomous molecular computer for logical control of gene expression. *Nature* **2004,** *429* (6990), 423-9.

95. Duose, D. Y.; Schweller, R. M.; Hittelman, W. N.; Diehl, M. R., Multiplexed and reiterative fluorescence labeling via DNA circuitry. *Bioconjug Chem* **2010,** *21* (12), 2327-31.

96. Venkataraman, S.; Dirks, R. M.; Ueda, C. T.; Pierce, N. A., Selective cell death mediated by small conditional RNAs. *Proc Natl Acad Sci U S A* **2010,** *107* (39), 16777-82.

97. Feng, B.; Frykholm, K.; Norden, B.; Westerlund, F., DNA strand exchange catalyzed by molecular crowding in PEG solutions. *Chem Commun (Camb)* **2010,** *46* (43), 8231-3.

98. Moll, J. R.; Ruvinov, S. B.; Pastan, I.; Vinson, C., Designed heterodimerizing leucine zippers with a ranger of pIs and stabilities up to 10(-15) M. *Protein Sci* **2001,** *10* (3), 649-55.

99. Rouillard, J. M.; Zuker, M.; Gulari, E., OligoArray 2.0: design of oligonucleotide probes for DNA microarrays using a thermodynamic approach. *Nucleic Acids Res* **2003,** *31* (12), 3057-62.

100. Lilley, D. M. J., Structures of helical junctions in nucleic acids. *Q Rev Biophys* **2000,** *33* (2), 109-59.

101. Karymov, M.; Daniel, D.; Sankey, O. F.; Lyubchenko, Y. L., Holliday junction dynamics and branch migration: single-molecule analysis. *Proc Natl Acad Sci U S A* **2005,** *102* (23), 8186-91.

102. Heilemann, M.; van de Linde, S.; Schuttpelz, M.; Kasper, R.; Seefeldt, B.; Mukherjee, A.; Tinnefeld, P.; Sauer, M., Subdiffraction-resolution fluorescence imaging with conventional fluorescent probes. *Angew Chem Int Ed Engl* **2008,** *47* (33), 6172-6.

103. Rothemund, P. W., Folding DNA to create nanoscale shapes and patterns. *Nature* **2006,** *440* (7082), 297-302.

104. Li, B.; Ellington, A. D.; Chen, X., Rational, modular adaptation of enzyme-free DNA circuits to multiple detection methods. *Nucleic Acids Res* **2011,** *39* (16), e110.

105. Modi, S.; M, G. S.; Goswami, D.; Gupta, G. D.; Mayor, S.; Krishnan, Y., A DNA nanomachine that maps spatial and temporal pH changes inside living cells. *Nat Nanotechnol* **2009,** *4* (5), 325-30.

106. Genot, A. J.; Bath, J.; Turberfield, A. J., Reversible logic circuits made of DNA. *J Am Chem Soc* **2011,** *133* (50), 20080-3.

107. Douglas, S. M.; Bachelet, I.; Church, G. M., A logic-gated nanorobot for targeted transport of molecular payloads. *Science* **2012,** *335* (6070), 831-4.

108. Schieker, M.; Pautke, C.; Haasters, F.; Schieker, J.; Docheva, D.; Bocker, W.; Guelkan, H.; Neth, P.; Jochum, M.; Mutschler, W., Human mesenchymal stem cells at the single-cell level: simultaneous seven-colour immunofluorescence. *J Anat* **2007,** *210* (5), 592-9.

109. Constantinou, P.; Dacosta, R. S.; Wilson, B. C., Extending immunofluorescence detection limits in whole paraffin-embedded formalin fixed tissues using hyperspectral confocal fluorescence imaging. *J Microsc* **2009,** *234* (2), 137-46.

110. Zimmermann, T., Spectral imaging and linear unmixing in light microscopy. *Adv Biochem Eng Biotechnol* **2005,** *95*, 245-65.

111. Duose, D. Y.; Schweller, R. M.; Zimak, J.; Rogers, A. R.; Hittelman, W. N.; Diehl, M. R., Configuring robust DNA strand displacement reactions for in situ molecular analyses. *Nucleic Acids Res* **2011**.

112. Uhlen, M.; Oksvold, P.; Fagerberg, L.; Lundberg, E.; Jonasson, K.; Forsberg, M.; Zwahlen, M.; Kampf, C.; Wester, K.; Hober, S.; Wernerus, H.; Bjorling, L.; Ponten, F., Towards a knowledge-based Human Protein Atlas. *Nat Biotechnol* **2010,** *28* (12), 1248-50.

113. Taylor, I. W.; Linding, R.; Warde-Farley, D.; Liu, Y.; Pesquita, C.; Faria, D.; Bull, S.; Pawson, T.; Morris, Q.; Wrana, J. L., Dynamic modularity in protein interaction networks predicts breast cancer outcome. *Nat Biotechnol* **2009,** *27* (2), 199-204.

114. Tulchin, N.; Chambon, M.; Juan, G.; Dikman, S.; Strauchen, J.; Ornstein, L.; Billack, B.; Woods, N. T.; Monteiro, A. N., BRCA1 protein and nucleolin colocalize in breast carcinoma tissue and cancer cell lines. *Am J Pathol* **2010,** *176* (3), 1203-14.

115. Zhou, M.; Veenstra, T. D., Proteomic analysis of protein complexes. *Proteomics* **2007,** *7* (16), 2688-97.

116. Magliery, T. J.; Wilson, C. G.; Pan, W.; Mishler, D.; Ghosh, I.; Hamilton, A. D.; Regan, L., Detecting protein-protein interactions with a green fluorescent protein fragment reassembly trap: scope and mechanism. *J Am Chem Soc* **2005,** *127* (1), 146-57.

117. Truong, K.; Ikura, M., The use of FRET imaging microscopy to detect protein-protein interactions and protein conformational changes in vivo. *Curr Opin Struct Biol* **2001,** *11* (5), 573-8.

118. Zuker, M., Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **2003,** *31* (13), 3406-15.

119. Dirks, R. M.; Bois, J. S.; Schaeffer, J. M.; Winfree, E.; Pierce, N. A., Thermodynamic analysis of interacting nucleic acid strands. *Siam Review* **2007,** *49* (1), 65-88.

120. Zadeh, J. N.; Steenberg, C. D.; Bois, J. S.; Wolfe, B. R.; Pierce, M. B.; Khan, A. R.; Dirks, R. M.; Pierce, N. A., NUPACK: Analysis and design of nucleic acid systems. *J Comput Chem* **2011,** *32* (1), 170-3.

121. Yang, F.; Moss, L. G.; Phillips, G. N., Jr., The molecular structure of green fluorescent protein. *Nat Biotechnol* **1996,** *14* (10), 1246-51.

122. Konig, T.; Skerra, A., Use of an albumin-binding domain for the selective immobilisation of recombinant capture antibody fragments on ELISA plates. *J Immunol Methods* **1998,** *218* (1-2), 73-83.

123. MacBeath, G.; Schreiber, S. L., Printing proteins as microarrays for high-throughput function determination. *Science* **2000,** *289* (5485), 1760-3.

124. Wu, G.; Datar, R. H.; Hansen, K. M.; Thundat, T.; Cote, R. J.; Majumdar, A., Bioassay of prostate-specific antigen (PSA) using microcantilevers. *Nat Biotechnol* **2001,** *19* (9), 856-60.

125. Camarero, J. A., Recent developments in the site-specific immobilization of proteins onto solid supports. *Biopolymers* **2008,** *90* (3), 450-8.

126. Rusmini, F.; Zhong, Z.; Feijen, J., Protein immobilization strategies for protein biochips. *Biomacromolecules* **2007,** *8* (6), 1775-89.

127. Zhu, H.; Snyder, M., Protein chip technology. *Curr Opin Chem Biol* **2003,** *7* (1), 55-63.

128. Niemeyer, C. M.; Boldt, L.; Ceyhan, B.; Blohm, D., DNA-Directed immobilization: efficient, reversible, and site-selective surface binding of proteins by means of covalent DNA-streptavidin conjugates. *Anal Biochem* **1999,** *268* (1), 54-63.

129. Wacker, R.; Niemeyer, C. M., DDI-microFIA--A readily configurable microarray-fluorescence immunoassay based on DNA-directed immobilization of proteins. *Chembiochem* **2004,** *5* (4), 453-9.

130. Wacker, R.; Schroder, H.; Niemeyer, C. M., Performance of antibody microarrays fabricated by either DNA-directed immobilization, direct spotting, or streptavidin-biotin attachment: a comparative study. *Anal Biochem* **2004,** *330* (2), 281-7.

131. Erben, C. M.; Goodman, R. P.; Turberfield, A. J., Single-molecule protein encapsulation in a rigid DNA cage. *Angew Chem Int Ed Engl* **2006,** *45* (44), 7414-7.

132. Jung, G. Y.; Stephanopoulos, G., A functional protein chip for pathway optimization and in vitro metabolic engineering. *Science* **2004,** *304* (5669), 428-31.

133. Niemeyer, C. M.; Koehler, J.; Wuerdemann, C., DNA-directed assembly of bienzymic complexes from in vivo biotinylated NAD(P)H:FMN oxidoreductase and luciferase. *Chembiochem* **2002,** *3* (2-3), 242-5.

134. Kukolka, F.; Niemeyer, C. M., Synthesis of fluorescent oligonucleotide--EYFP conjugate: towards supramolecular construction of semisynthetic biomolecular antennae. *Org Biomol Chem* **2004,** *2* (15), 2203-6.

135. Niemeyer, C. M.; Sano, T.; Smith, C. L.; Cantor, C. R., Oligonucleotide-directed self-assembly of proteins: semisynthetic DNA--streptavidin hybrid molecules as connectors for the generation of macroscopic arrays and the construction of supramolecular bioconjugates. *Nucleic Acids Res* **1994,** *22* (25), 5530-9.

136. Kwon, Y.; Coleman, M. A.; Camarero, J. A., Selective immobilization of proteins onto solid supports through split-intein-mediated protein trans-splicing. *Angew Chem Int Ed Engl* **2006,** *45* (11), 1726-9.

137. Lovrinovic, M.; Seidel, R.; Wacker, R.; Schroeder, H.; Seitz, O.; Engelhard, M.; Goody, R. S.; Niemeyer, C. M., Synthesis of protein-nucleic acid conjugates by expressed protein ligation. *Chem Commun (Camb)* **2003,** (7), 822-3.

138. Jongsma, M. A.; Litjens, R. H., Self-assembling protein arrays on DNA chips by auto-labeling fusion proteins with a single DNA address. *Proteomics* **2006,** *6* (9), 2650-5.

139. Duckworth, B. P.; Chen, Y.; Wollack, J. W.; Sham, Y.; Mueller, J. D.; Taton, T. A.; Distefano, M. D., A universal method for the preparation of covalent protein-DNA conjugates for use in creating protein nanostructures. *Angew Chem Int Ed Engl* **2007,** *46* (46), 8819-22.

140. Hoover, D. M.; Lubkowski, J., DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis. *Nucleic Acids Res* **2002,** *30* (10), e43.

141. Seeman, N. C., De novo design of sequences for nucleic acid structural engineering. *J Biomol Struct Dyn* **1990,** *8* (3), 573-81.

142. Meyer, D. E.; Chilkoti, A., Genetically encoded synthesis of protein-based polymers with precisely specified molecular weight and sequence by recursive directional ligation: examples from the elastin-like polypeptide system. *Biomacromolecules* **2002,** *3* (2), 357-67.

143. Zhang, K.; Diehl, M. R.; Tirrell, D. A., Artificial polypeptide scaffold for protein immobilization. *J Am Chem Soc* **2005,** *127* (29), 10136-7.

144. Diehl, M. R.; Zhang, K.; Lee, H. J.; Tirrell, D. A., Engineering cooperativity in biomotor-protein assemblies. *Science* **2006,** *311* (5766), 1468-71.

Appendix A

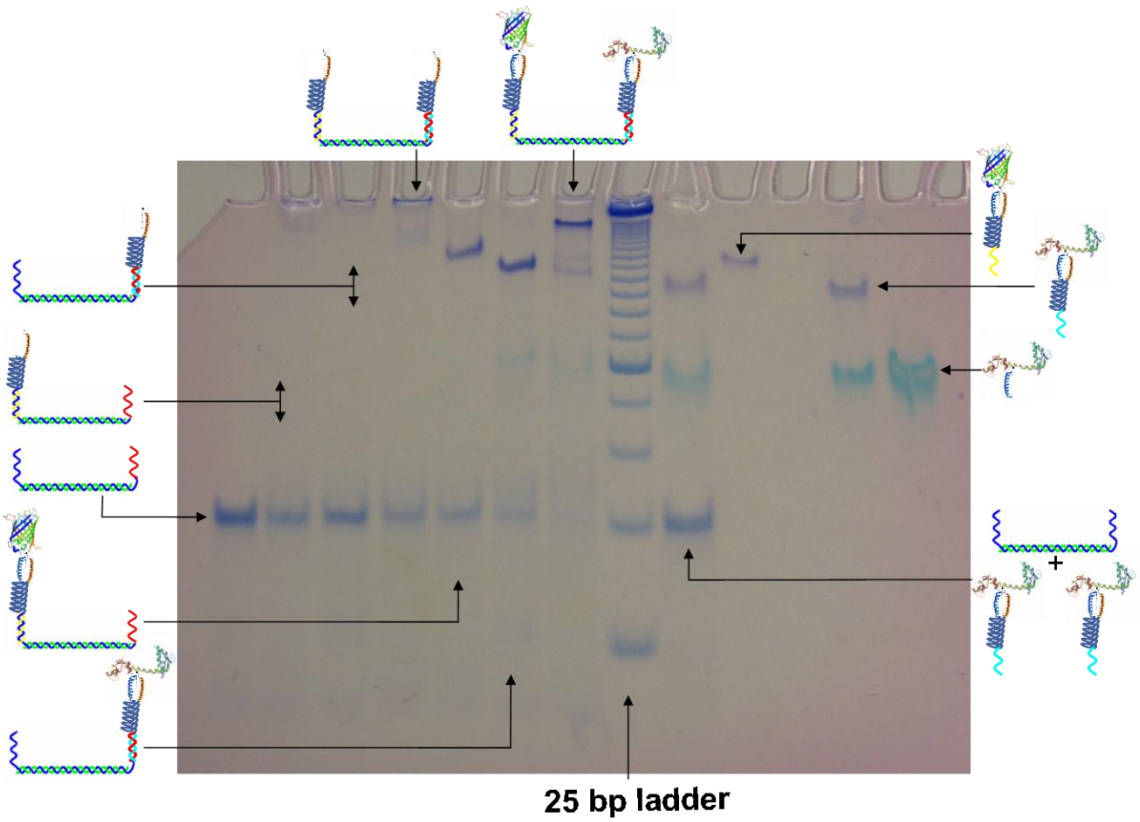

Figure A1. Full non-denaturing PAGE gel (Figure 7.7) displaying each of the components of the assembly individually as well as intermediate structures. The corresponding components of each lane are depicted at the periphery of the gel. From this analysis, we are able to adjust the volumes of individual protein stocks to compensate for errors in protein concentration measurements, in order to drive assembly formation more quantitatively.